

Ulrich Engelke | Barbara Engelke

ChatGPT

Mit KI in ein neues Zeitalter

Wie KI-Tools unser Leben und die Gesellschaft verändern

mitp


mitp

Ulrich Engelke
Barbara Engelke

ChatGPT

Mit KI in ein neues Zeitalter

Wie KI-Tools unser Leben und die Gesellschaft verändern

[image]

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISBN 978-3-7475-0759-9
1. Auflage 2023

www.mitp.de
E-Mail: mitp-verlag@sigloch.com
Telefon: +49 7953 / 7189 – 079
Telefax: +49 7953 / 7189 – 082

Lektorat: Janina Vervost
Sprachkorrektorat: Jürgen Benvenuti
Covergestaltung: Christian Kalkert
Bildnachweis: ©Zaleman \stock.adobe.com
Satz: mitp Verlags GmbH & Co. KG, Frechen, www.mitp.de
electronic **pub**lication: CPI books GmbH, Leck

Inhalt

Anmerkungen

Vorwort

E-Bikes sind böse! Aber sind sie es wirklich? Aus Sicht des Radsportlers ist die Sache klar: »Gift für die Beine und eine weitere Krücke für eine degenerierte Gesellschaft, die sich bald nicht mehr ohne künstliche Hilfsmittel fortbewegen kann.« Gott sei Dank gibt es nicht viele von diesen kurzsichtigen Sportlern, denn eigentlich hat das E-Bike viel Gutes gebracht: Die Menschen radeln dank der Unterstützung nun wieder, und zwar mehr und weiter als je zuvor!

Warum erzählen wir Ihnen das alles? Weil es Parallelen zu ChatGPT gibt, weil wir mehr als zehn Jahre nach dem E-Bike eine weitere *Krücke* bekommen haben: ChatGPT nämlich. Statt schlagender Radfahrer-Verbindungen verkünden nun Bildungsfunktionäre und die Wächter des Deutschunterrichtes den geistigen Niedergang des Abendlandes: Wenn wir ChatGPT das Schreiben überlassen und sogar Schüler das Werkzeug in der Schule nutzen, kann bald niemand mehr selbst formulieren und Texte schreiben. Schlimmer noch: Weil Denken und Sprache nun mal untrennbar sind, sind die intellektuelle Entwicklung und die geistige Reife künftiger Generationen in akuter Gefahr.

Echt jetzt? So schlimm? Haben wir eigentlich beim Aufkommen des Taschenrechners denselben Tanz aufgeführt? Und wie war es damals bei der Erfindung des Rades? Hat man damals auch aufgeschrien, wir würden jetzt das Gehen und das Tragen schwerer Dinge verlernen? Wohl wahr: Seit es das Auto gibt, können nur noch wenige Menschen reiten. Aber ist diese Entwicklung so fatal, wie man vor mehr als 100 Jahren vielleicht gedacht hatte?

ChatGPT ist seit dem 30. 11. 2022 für alle zugänglich und nutzbar. Es ist ein Faktum wie das Auto und der Taschenrechner. Wir können nur noch darauf reagieren, individuell und als Gesellschaft. Und versuchen, das Beste daraus zu machen. Dieses Buch soll Ihnen dabei helfen und Licht auf folgende Fragen werfen:

- Was ist künstliche Intelligenz?
- Wie entsteht sie und was kann sie leisten?
- Wie wird sie unser Leben und unsere Gesellschaft verändern?

Wir haben uns Mühe gegeben, diese Fragen und weitere umfassend und verständlich zu beantworten. Und hoffen, dass unsere Antworten Sie weiterbringen.

1KI und Machine Learning: Entwicklung und Technologien

1.1Natürliche Intelligenz und ihre künstliche Schwester

Wer sich mit künstlicher Intelligenz (KI oder engl. AI für *artificial intelligence*) beschäftigt, muss zuerst klären, was natürliche Intelligenz eigentlich ist. *Intelligenz* ist ein ziemlich schwer zu definierendes Konzept. Wissenschaftliche Erklärungsversuche kommen aus der Psychologie und Philosophie, der Neurologie und weiteren Disziplinen, die basierend auf ihrem Forschungsgegenstand jeweils andere Aspekte in den Vordergrund stellen. Eine zusammenhängende Erklärung oder ein tiefgehendes Verständnis, warum unser Gehirn zu wirklich außerordentlichen Leistungen fähig ist, gibt es noch nicht.

Versucht man sich an einer allgemeingültigen Definition, die quer durch alle Fachrichtungen wenigstens ein Kopfnicken hervorrufen soll, könnte man Intelligenz als die Fähigkeit bezeichnen, Wissen zu erwerben und dieses Wissen anzuwenden, um damit beliebige Probleme zu lösen. In einem weiteren Schritt entsteht daraus idealerweise neues Wissen und es können neue Erkenntnisse und Konzepte formuliert werden. Intelligenz ist untrennbar mit dem Begriff des *Transfers* verbunden. Es geht dabei um Anpassungsfähigkeit, das Lernen aus Erfahrungen und das Anwenden von erworbenem Wissen in stetig variierenden Kontexten. Die Entwicklung von Intelligenztests zur Bewertung und Messung menschlicher Intelligenz war von Anfang an ein fragwürdiges Unterfangen, das nicht zu einem besseren Verständnis des Phänomens geführt hat.

Künstliche Intelligenz war und ist ganz allgemein gesprochen der Versuch, menschliche Intelligenz in Maschinenform nachzubilden und irgendwann sogar zu übertreffen. Maschinen im Allgemeinen und Computer im Besonderen sollten so gestaltet werden, dass sie menschenähnliche Fähigkeiten erhalten, um Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern. Künstliche Intelligenz als umfassender Ansatz wurde in viele einzelne Teilbereiche zerlegt, von Computer Vision (Bilder und Videos interpretieren, um daraus Informationen zu gewinnen) bis zur Robotik (Entwicklung von Maschinen, die eine Vielzahl von Aufgaben autonom ausführen können). Die natürliche Sprachverarbeitung (Natural Language Processing, NLP), also das Verstehen und aktive Verwenden menschlicher Sprache für Anwendungen wie Chatbots, Übersetzungsdienste und Sprachassistenten,

entwickelte sich schnell zu einem der anspruchsvollsten Forschungszweige in der Informatik.

Künstliche Intelligenz ist bis heute keine Nachbildung *allgemeiner* menschlicher Intelligenz in Maschinenform. Die Entwicklung konkreter Anwendungen konzentriert sich immer auf eine *spezialisierte* Intelligenz für ganz bestimmte Aufgaben. Ein Schachcomputer kann die besten Spieler der Welt auf dem Schachbrett mattsetzen, kann außerhalb des Spielfeldes aber kein einfaches Gespräch führen und eine Maus nicht von einem Elefanten unterscheiden.

Heute ist die künstliche Repräsentation menschlicher Intelligenz ein ziemlich buntes und multidisziplinäres Unterfangen, das ständig wächst und sich in alle Richtungen in rasendem Tempo weiterentwickelt: Chatbots sprechen wie Menschen mit uns, Computerprogramme erkennen den Inhalt von Bildern, produzieren auf Wunsch auch neue Bilder oder Videos in täuschend echter Qualität, sie erkennen und bewerten komplexe Verkehrssituationen in Echtzeit und können Fahrzeuge autonom steuern. Diesen wirklich beeindruckenden Fortschritten in vielen Bereichen, die menschliche Fähigkeiten teilweise um ein Vielfaches überschreiten, stehen andere Aspekte menschlicher Intelligenz gegenüber, die auf der Maschine offenbar schwer zu replizieren sind: Emotion, Empathie und interdisziplinäres Denken.

1.2Ursprünge der KI und historische Entwicklung

Die Geschichte der künstlichen Intelligenz ist ebenso lang wie faszinierend. Erste Konzepte und Visionen stammen von Heron von Alexandria, einem griechischen Mathematiker und Ingenieur, der unter dem Titel *Automata* das erste *Buch der Maschinen* veröffentlichte und einen Weihwasserautomaten entwickelte. Die Jahrtausende alte Idee, künstliches Leben zu erschaffen, griff ein französischer Ingenieur und Erfinder namens Jacques de Vaucanson auf und entwickelte 1738 die *Mechanische Ente* und weitere Automaten. Aber erst das 20. Jahrhundert gilt als Geburtsstunde der modernen KI. Pionier in den 1930er- und 1940er-Jahren war Alan Turing mit seiner *Turing-Maschine*, die kein physisches Gerät, sondern ein theoretisches Konzept zur Theorie der Berechenbarkeit darstellte. Sie hat zur Entwicklung von Programmiersprachen, zur Theorie der Automaten und zur Erkenntnis der Grenzen der Computermathematik beigetragen und Grundlagen für das Verständnis algorithmischer Prozesse gelegt. Von ihm stammt auch der bekannte *Turing-Test*, der ab 1950 zu einem zentralen Diskussionsthema im Bereich KI wurde: Ein Computer besteht den

Turing-Test, wenn ein menschlicher *Richter* nicht in der Lage ist zu entscheiden, ob er mit einem Menschen oder einem Computer kommuniziert, basierend allein auf den gegebenen Antworten.

Der Begriff *künstliche Intelligenz* wurde 1955 von John McCarthy geprägt und im Rahmen des Dartmouth Meetings im Jahr 1956 vorgestellt. Dieser Kongress gilt als die Geburtsstunde der KI als eigenständige Teildisziplin der Informatik. Die ersten KI-Programme wurden mit großer Euphorie entwickelt, darunter das Schachprogramm von Claude Shannon und das Programm *Logic Theorist* von Allen Newell und Herbert A. Simon. In den 1960ern und 1970ern gab es einen großen Optimismus in der KI-Forschung, mit respektablen Fortschritten in Bereichen wie maschinellem Lernen und der Sprachverarbeitung. »ELIZA« wurde in den 1960er-Jahren von Joseph Weizenbaum am Massachusetts Institute of Technology (MIT) entwickelt und war eines der ersten Computerprogramme, das in der Lage war, textbasierte Konversationen mit Menschen zu führen. Auf Grundlage des DOCTOR-Skripts ahmte es die Gesprächsführung eines Therapeuten nach. ELIZA verwendet Muster und Schlüsselwörter, um die Eingabe eines Benutzers zu analysieren und passende vordefinierte Antworten auszuwählen. Viele Benutzer hatten das Gefühl, mit einem menschlichen Therapeuten und nicht mit einem Computer zu sprechen.

Allerdings traten in den späten 1970ern auch die Grenzen der KI-Techniken zutage und führten zum ersten *KI-Winter*. In dieser Periode erlahmte das allgemeine Interesse am Forschungsgebiet und die Finanzierung wurde stark zurückgefahren. In den 1980ern kamen die sogenannten *Expertensysteme* auf den Markt und sorgten für eine Wiederbelebung der KI. Diese waren darauf ausgelegt, menschliches Fachwissen in einem spezifischen, stark eingegrenzten Bereich zu sammeln und das Wissen vieler Experten darin zu bündeln, um Beratung und Entscheidungsunterstützung anzubieten. Ein prominentes Beispiel eines Expertensystems ist *MYCIN*, ein System, das Ärzte bei der Diagnose von bakteriellen Infektionen unterstützt und Therapieempfehlungen für Antibiotika gibt. Weitere Systeme wie *PUFF* halfen Medizinern bei der Interpretation von Lungenfunktionsdaten, ein anderes unterstütze Ingenieure bei der Entwicklung von Schaltungen (*CADET*: Computer Aided Design of Electric Circuits). Alle diese Systeme markierten einen signifikanten Fortschritt in der KI-Forschung und schufen überaus nützliche Anwendungen für eng abgegrenzte Bereiche. Obwohl ihre Fähigkeiten im Vergleich zu heutigen KI-Systemen sehr begrenzt waren, legten sie den Grundstein für viele heutige KI-Anwendungen und maschinelle Lernsysteme.

Auf den ersten KI-Winter folgte der zweite, als gegen Ende des Jahrzehnts die anfängliche Begeisterung wieder verflogen war. Ein

Rückgang des allgemeinen Interesses an einem Forschungsgebiet führt immer zu einer Reduzierung der finanziellen Mittel und diese Ausdünnung der Forschungsgelder führte unweigerlich zum zweiten (und letzten) KI-Winter. Die Grundlagenforschung in den 1990ern fokussierte sich immer weiter auf datengetriebene, statistische Ansätze, besonders in der Sprachverarbeitung und beim maschinellen Lernen. In den 2000ern beschleunigten sich die Entwicklungen dann weiter. Die Verfügbarkeit von immer leistungsstärkerer Hardware und die exponentielle Zunahme an verfügbaren Daten läuteten nun das Zeitalter von *Big Data* ein und sorgten für große Fortschritte im Bereich des maschinellen Lernens.

Das 21. Jahrhundert bedeutet für die KI also gleichermaßen Renaissance wie endgültiger Durchbruch. Deep Learning sowie neuronale Netzwerke mit vielen Schichten und tiefen Architekturen revolutionierten Bereiche wie Bild- und Spracherkennung. KI-Systeme wie IBMs *Watson* und *AlphaGo* von DeepMind zeigten beeindruckende Fähigkeiten in der Lösung immer komplexerer Aufgaben. Bis heute hat die KI-Forschung eine beeindruckende Reise hinter sich, von ersten Visionen über die theoretischen Anfänge bis hin zu realen Anwendungen, die unsere moderne Welt immer tiefer prägen. Noch immer gibt es zahlreiche ungelöste Fragen und Herausforderungen, aber die Fortschritte der letzten Jahrzehnte sind erstaunlich.

1.3KI-Sprachassistenten und Gründung von OpenAI

Ich erinnere mich an eine Szene aus der Serie *Raumschiff Enterprise*, die mich als Kind sehr beeindruckt hat: Ein Außerirdischer erhält Zugriff auf den Bordcomputer des Raumschiffes und eröffnet seine Befehlseingabe ganz selbstverständlich mit »Hallo Computer«. Ein etwas verdutzter, aber hilfsbereiter Mr. Spock händigt ihm daraufhin eine Tastatur aus, die der andere mit verächtlicher Miene entgegennimmt und mit »Wie rückständig!« kommentiert. Was in den 1970ern noch reine Utopie war, ist heute längst Realität: Siri, Alexa und ähnliche Sprachassistenten gehören auf dem Smartphone zum Alltag und sind dort längst keine Sensation mehr. Vor der Einführung von Siri im Jahre 2011 als Teil des iPhone 4S waren KI-Interaktionen mit gesprochener Sprache für den durchschnittlichen Verbraucher auf Science-Fiction-Filme beschränkt. Auf Apples Siri folgte Amazons Alexa und der Google Assistant. Die Sprachassistenten brachten KI in die Taschen und Wohnzimmer von zahllosen Menschen. Aus dem ehrgeizigen Konzept wurde im Handumdrehen ein ganz alltägliches Werkzeug.

Die Sprachassistenten der großen Internet-Konzerne führten mit der Kommerzialisierung zu erheblichen Investitionen in Forschung und Entwicklung von Spracherkennungstechnologien. Fortschritte in der Spracherkennung haben die Popularität von Sprachassistenten noch weiter gesteigert und die anfängliche Fehlerquote in der Erkennung natürlicher gesprochener Sprache ist in den letzten Jahren drastisch gesunken. Hinzu kommen Fortschritte im Natural Language Processing (NLP), weil Assistenten auch den Kontext und die Bedeutung der Anfragen verstehen müssen. Dieses konkrete kommerzielle Interesse hat die Entwicklung in den Bereichen NLP und Natural Language Understanding (NLU) enorm vorangetrieben.

Ein Meilenstein in der Entwicklung der KI war die Gründung von OpenAI im Dezember 2015. Die Organisation wollte sich der Forschung im Bereich künstlicher Intelligenz widmen. Der im ursprünglichen Gründungsbrief dargelegte Hauptzweck formulierte, dass künstliche allgemeine Intelligenz (Artificial General Intelligence, AGI) der gesamten Menschheit zugutekommen sollte. OpenAI wollte selbst Anwendungen entwickeln und damit sicherstellen, dass KI auch von anderen auf eine Weise entwickelt wird, die der Menschheit nützt. Zu den Gründern von OpenAI gehörten Tech-Unternehmer und -Investoren wie Elon Musk und Sam Altman, der später CEO wurde. OpenAI hatte sich von Beginn an darauf festgelegt, seine Forschung öffentlich zugänglich zu machen, später wurde diese vollständige Offenheit wegen Sicherheits- und Wettbewerbsbedenken aber immer weiter eingeschränkt.

Elon Musk war einer der Hauptunterstützer von OpenAI, hatte aber keine Kontrolle über die Aktivitäten der Organisation, die als unabhängige Einheit agierte. Musk war natürlich Vorstandsmitglied und wollte der Nachrichtenwebsite Semafor zufolge im Jahr 2018 die Kontrolle von OpenAI übernehmen. CEO Sam Altman und andere Gründer hätten seine Ambitionen aber abgelehnt, was zum Ausstieg von Musk wegen angeblicher Interessenkonflikte führte. Die Gründung von OpenAI war neben gemeinnützigen und geschäftlichen Interessen nicht zuletzt auch Ausdruck einer wachsenden Besorgnis in der Tech-Gemeinschaft über die potenziellen Risiken von KI und insbesondere von AGI. Durch ihre Bemühungen und Forschungen hoffte die Organisation damals, den Weg für eine sichere und für alle vorteilhafte Entwicklung von fortschrittlichen KI-Technologien zu ebnen.

1.4Machine Learning (ML) und Deep Learning (DL)

Machine Learning oder maschinelles Lernen ist ganz allgemein

gesprochen der Ansatz, Maschinen so zu programmieren, dass sie, statt vorgegebene Anweisungen zu befolgen, aus Daten lernen können, ähnlich wie Menschen aus ihren Erfahrungen lernen. Ziel ist, dass die Maschinen ihre Leistung bei der Lösung bestimmter Aufgaben über die Zeit hinweg durch Erfahrung (also weiteres Lernen) immer weiter verbessern.

Machine Learning (ML) und Deep Learning (DL) sind Teilgebiete der künstlichen Intelligenz, die sich aber in Konzept, Architektur und Anwendung stark unterscheiden. Machine Learning beschreibt den konzeptuellen Ansatz, bei dem Computer die Fähigkeit erlernen, Aufgaben ohne explizite Programmierung auszuführen. ML-Algorithmen nutzen vorgegebene Daten, um Vorhersagen oder Entscheidungen zu treffen, ohne explizite Programmanweisungen, wie diese Entscheidungen getroffen werden sollen. Wenn ein KI-Programm Bilder von Stühlen und Sesseln unterscheiden soll, werden im Algorithmus keine Kriterien für Stühle oder Sessel hinterlegt. Stattdessen werden dem Programm gelabelte Trainingsdaten vorgelegt, also Bilder von Stühlen und Sesseln gezeigt, und für jedes Bild wird der korrekte Typ genannt. Der ML-Algorithmen muss (und kann) nun selbstständig geeignete Kriterien zur Unterscheidung finden und gewichten. Nach den Trainingsdaten kommen weitere, bisher unbekannte Testdaten, mit denen geprüft wird, ob das System bereits praxistauglich ist oder weiteres Training benötigt.

Deep Learning (DL) ist ein Teilbereich von ML, der meistens tief verschachtelte neuronale Netzwerkarchitekturen verwendet. Diese *tiefen* Modelle können aus sehr vielen Schichten bestehen, weshalb man sie als *Deep Learning* bezeichnet. DL-Modelle sind komplexe Netzwerke, die sehr große Mengen an Daten benötigen, um angemessen trainiert zu werden. ML-Modelle können mit deutlich weniger Daten auskommen als DL-Modelle. Wegen der Datenmenge und der höheren Komplexität der Netzwerke erfordern DL-Modelle häufig spezialisierte Hardware wie Grafikprozessoreinheiten (GPUs) oder Tensor Processing Units (TPUs) für das Training, was enorm rechen- und kostenintensiv ist. Einfache ML-Modelle sind deutlich anspruchsloser und erfordern keine besondere Hardware.

Deep-Learning-Modelle, die auf großen Datenmengen beruhen, sind zwar sehr leistungsfähig, weil es aber schwierig bis unmöglich ist, die genaue Funktionsweise der einzelnen Neuronenschichten nachzuvollziehen, sind die Ergebnisse wenig bis gar nicht interpretierbar und werden oft als *Black Boxes* empfunden. Machine Learning wird in eher einfachen Anwendungen der Vorhersageanalyse, für Empfehlungssysteme bis hin zur Mustererkennung eingesetzt. Deep Learning hingegen ist bei sehr großen Mengen unstrukturierter Daten im Vorteil und entwickelt das volle Potenzial bei Aufgaben wie der Bild-

und Spracherkennung. ML nutzt eine Vielzahl von Algorithmen wie lineare Regression, Entscheidungsbäume und Support Vector Machines, DL konzentriert sich hauptsächlich auf neuronale Netzwerkarchitekturen und Transformer, wie sie auch bei ChatGPT zum Einsatz kommen.

Deep Learning ist also so etwas wie das Sahnehäubchen und die Weiterentwicklung von Machine Learning. Es setzt auf komplexe neuronale Netzwerke und deren tief verschachtelte Architekturen. Praktisch alle neueren Leuchtturm-Anwendungen der KI basieren auf Deep-Learning-Algorithmen und profitieren von der großen Menge an verfügbaren Daten und einer immer leistungsfähigeren Hardware mit exponentiell wachsender Rechenleistung. Wussten Sie das schon? Nach heutigen Standards betrachtet, hat ein aktuelles Smartphone schon deutlich mehr Rechenleistung als die Computer der ersten Apollo-Mission!

1.5Schlüsseltechnologien und Methoden des Machine Learnings

Künstliche Intelligenz umfasst eine Reihe von Schlüsseltechnologien für spezifische Anwendungen, die je nach Problem und Art der vorliegenden Daten eingesetzt werden. Beim Machine Learning werden vor allem drei unterschiedliche Formen des Lernens eingesetzt: das Supervised Learning, das Unsupervised Learning und das Reinforcement Learning.

Man kann ein Modell anhand von gelabelten Daten trainieren, in unserem Beispiel Bilder von Stühlen und Sesseln, die vorher *gelabelt*, also gekennzeichnet und der jeweiligen Klasse zugeordnet wurden. Das Verfahren zum Training des Modells nennt sich überwachtes Lernen oder **Supervised Learning**. Typische Anwendungen von überwachtem Lernen sind E-Mail-Spam-Filter: Ein Modell wird trainiert, um zu erkennen, ob eine E-Mail Spam ist oder nicht. Das Training erfolgt auf einem Datensatz, in dem vorhandene E-Mails bereits als *Spam* oder *Nicht-Spam* gekennzeichnet sind. Systemanalytiker bezeichnen dieses Verfahren als *Klassifikation*. Im Gegensatz dazu eröffnen sich mit der sogenannten *Regression* weitere Möglichkeiten, wie sie z. B. in der Vorhersage von Immobilienpreisen durch eine KI verwendet werden: Ein Modell wird trainiert, um den Verkaufspreis von Häusern, basierend auf verschiedenen Merkmalen wie Größe, Lage und Anzahl der Zimmer, vorherzusagen. Der Trainingsdatensatz enthält historische Daten von verkauften Häusern mit allen Merkmalen und den tatsächlich erzielten Verkaufspreisen.

Wenn die Daten zum Trainieren von Modellen ohne Labels

auskommen müssen, sprechen wir vom unüberwachten Lernen bzw. **Unsupervised Learning**. Das kommt immer dann zum Einsatz, wenn es zu viele Parameter gibt und gleichzeitig die Relevanz einzelner Parameter noch nicht eingeschätzt werden kann. Dass man auf die Intelligenz eines Menschen eher über dessen Schulnoten als über die Schuhgröße rückschließen kann, steht sicherlich außer Frage. Aber an welchen Parametern erkennt der Online-Händler die Besucher seiner Website mit dem größten Umsatz-Potenzial?

Beim unüberwachten Lernen werden Modelle anhand von nicht gelabelten Daten trainiert. Das Hauptziel besteht darin, Muster, Strukturen oder Zusammenhänge in den Daten zu entdecken, ohne dass vorher festgelegte Labels oder Kategorien vorhanden sind. An konkreten Beispielen wird das deutlicher: Ein Handelsunternehmen möchte seine Kunden basierend auf ihrem Kaufverhalten in verschiedene Gruppen einteilen, ohne vorher zu wissen, welche Kategorien dafür eigentlich relevant sind. Ein **Clustering-Algorithmus** wie *k-Means* kann verwendet werden, um die Kunden in verschiedene Gruppen bzw. Cluster zu unterteilen, ohne dass vorher Kategorien für die Einteilung festgelegt wurden. Clustering ist für alle Arten von Segmentierungen ein sinnvolles Verfahren.

Für die Analyse des Warenkorbes werden **Assoziationsregeln** verwendet: Ein Einzelhändler möchte herausfinden, welche Produkte häufig zusammen gekauft werden, um effektive Produktbündel oder Werbeaktionen zu erstellen. Ein Algorithmus wie *Apriori* kann verwendet werden, um solche Assoziationsregeln in vorliegenden Transaktionsdaten zu finden. Bei diesen Beispielen zum Clustering und Assoziationsregel-Lernen gibt es keine vorgegebenen Labels oder Kategorien. Das Modell versucht stattdessen, die zugrunde liegende Struktur der Daten zu erkennen und daraus Erkenntnisse zu gewinnen.

Die dritte Form des Lernens ist das sogenannte **Reinforcement Learning** oder verstärkendes Lernen. Modelle lernen durch Belohnungen, basierend auf den getroffenen Entscheidungen. Das Konzept des verstärkenden Lernens hat Parallelen zur positiven und negativen Verstärkung in der Verhaltenspsychologie, die auch in Erziehungskontexten angewendet wird. Tatsächlich stammen viele Prinzipien des verstärkenden Lernens in der KI aus der Verhaltenspsychologie. In der Erziehung von Kindern bedeutet Verstärkung, dass bestimmte Reaktionen oder Verhaltensweisen des Kindes durch Belohnungen (positive Verstärkung) oder unangenehme Reize (negative Verstärkung) gestärkt werden, um das gewünschte Verhalten zu fördern. Maschinelles verstärkendes Lernen wird oft in der Robotik und in der Spiele-KI eingesetzt.

AlphaGo ist ein Programm, das von DeepMind entwickelt wurde, um

das Brettspiel *Go* zu spielen. Es hat das Prinzip des verstärkenden Lernens verwendet, um sich selbst zu trainieren und schließlich den menschlichen Weltmeister im Go zu schlagen. Oder wir nehmen ein Beispiel aus der Roboter-Navigation: Ein Roboter wird in einer unbekannten Umgebung platziert und muss lernen, von einem Punkt zum anderen zu navigieren, ohne gegen Hindernisse zu stoßen. Durch verstärkendes Lernen kann der Roboter lernen, welche Aktionen ihn sicher zu seinem Ziel führen und welche zu Kollisionen oder anderen negativen Ergebnissen führen. Zuletzt noch ein Beispiel aus der Finanzwelt bzw. dem Aktienhandel: Ein Algorithmus kann trainiert werden, um Kauf- und Verkaufsentscheidungen auf dem Aktienmarkt zu treffen, wobei das Ziel (die Belohnung) ist, den finanziellen Gewinn zu maximieren. Durch Interaktion mit Marktinformationen und -preisen kann der Algorithmus über die Zeit lernen, welche Handelsstrategien am rentabelsten sind.

In allen diesen Beispielen interagiert ein Agent (z. B. das AlphaGo-Programm, der Roboter oder der Handelsalgorithmus) mit seiner Umgebung, trifft Entscheidungen und erhält Feedback in Form von Belohnung (Erfolg) oder Nicht-Belohnung (Misserfolg). Der Algorithmus kann sich entsprechend anpassen, um die Wahrscheinlichkeit von Belohnungen über die Zeit zu maximieren.

1.6Schlüsseltechnologien und Methoden des Deep Learnings

Deutlich komplexer sind die Schlüsseltechnologien und Methoden des Deep Learnings. Hier werden in der Bildverarbeitung vorwiegend *Konvolutionale neuronale Netzwerke*, für sequenzielle Daten wie Text oder Zeitreihen *Rekurrente neuronale Netzwerke* oder *Transformer-Architekturen* wie bei GPT eingesetzt.

Transformer-Architekturen haben den Bereich des Natural Language Processing (NLP) revolutioniert. Die hier vorgestellten Konzepte sind sehr komplex, ermöglichen aber einen Einblick in die spannendsten Anwendungen moderner KI, die nur durch Deep Learning, große Datenmengen und mit ausgefeilten Technologien erst möglich wurden. Wem die folgenden Erklärungen zu technisch sind, darf die nächsten Abschnitte natürlich gern überspringen.

Konvolutionale neuronale Netzwerke (Convolutional Neural Networks, CNNs oder ConvNets) sind eine Kategorie von tiefen neuronalen Netzwerken, die speziell für Aufgaben wie die Bilderkennung entwickelt wurden. Die Grundidee von CNNs besteht darin, die

räumlichen Hierarchien in Daten zu erkennen. CNNs sind so strukturiert, dass sie automatisch und adaptiv Merkmale aus Bildern lernen, verallgemeinern und wieder kombinieren können. In der ersten Schicht des Netzwerks (Convolutional Layer) werden Filter verwendet, um lokale Merkmale eines Bildes zu erkennen und eine sogenannte *Feature-Map* zu produzieren. Nach vielen weiteren Schichten werden die Daten in eine oder mehrere vollständig vernetzte Schichten geführt, um das endgültige Ergebnis (z. B. eine Klassifikation) zu erhalten.

Stellen Sie sich das Bild einer Katze im Garten vor: Zuerst werden einfache Merkmale wie Kanten und Texturen erkannt, in den tieferen Schichten folgen komplexere Merkmale wie Augen, Nasen und Ohren. Noch tiefer könnten weitere Objekte wie Blumen oder Bäume erkannt werden. Erst die unteren *Pooling*-Schichten können diese Merkmale dann verallgemeinern und in den *vollständig vernetzten Schichten* zu einer Schlussfolgerung gelangen, dass es sich bei dem Bild um eine »Katze im Garten« handelt. Dieser Prozess des Erkennens von Merkmalen von einfach bis komplex und deren Zusammenführung ermöglicht es CNNs, Bilder mit hoher Genauigkeit zu klassifizieren, um dann in der Praxis bildbezogene Aufgaben von der Gesichtserkennung bis zur medizinischen Bildanalyse zu lösen.

Rekurrente neuronale Netzwerke (RNNs) sind eine Klasse von neuronalen Netzwerken, die speziell für die Verarbeitung sequenzieller Daten entwickelt wurden. Sequenzielle Daten sind Informationen, die in geordneter Abfolge auftreten und bei denen die Reihenfolge für das Verständnis oder die Analyse wichtig ist. Dafür sorgen Schleifen, die es RNNs ermöglichen, Informationen über vorherige Schritte im Netzwerk zu behalten. Diese werden als *versteckte Zustände* gespeichert. Das macht RNNs besonders geeignet für Aufgaben, bei denen zeitliche Abhängigkeiten oder Sequenzen wichtig sind, wie in der Text- oder Spracherkennung. In der Praxis sind sie mittlerweile in der Lage, sehr komplexe Sequenzmuster zu erkennen, was sie für eine Vielzahl von Anwendungen empfiehlt, von der maschinellen Übersetzung über die Spracherkennung bis hin zur Komposition von Musikstücken.

Stellen Sie sich vor, Sie möchten den Satz »Als ich im Wald spazieren ging, hörte ich das Zwitschern eines __ .« vervollständigen. Für die Lücke wäre das Wort »Vogel« in der flektierten Form »Vogels« eine sinnvolle Lösung. Um zu diesem Schluss zu kommen, muss ein KI-Modell die Information über den Ort »Wald« und den Laut »Zwitschern« *erinnern*. Ein RNN verarbeitet den Satz Wort für Wort und aktualisiert seinen internen Zustand bei jedem einzelnen Wort neu. Wenn es zur Lücke kommt (und bei der Sprachgenerierung ist jedes nächste Wort eine Lücke), würde das RNN seinen aktuellen versteckten Zustand und die Information über den Wald und das Zwitschern verwenden, um das Wort

»Vogel« vorherzusagen und die korrekte Form »Vogels« zu finden.

Zuletzt wollen wir uns mit den **Transformer-Architekturen** beschäftigen, die zumindest im Bereich der Sprachverarbeitung aktuell die Speerspitze in den KI-Technologien darstellen. Transformer-Architekturen sind ein vergleichsweise neues Modell in der Deep-Learning-Landschaft, das speziell für Aufgaben in der Sequenz-zu-Sequenz-Verarbeitung, insbesondere in der Sprachverarbeitung, entwickelt wurde. Die Schlüsselkomponente eines Transformers ist ein *Self-Attention-Mechanismus*, der es diesen Modellen ermöglicht, verschiedene Positionen innerhalb einer Eingabesequenz in Abhängigkeit von ihrem Kontext zu gewichten.

Der Unterschied zu den vorher genannten RNNs macht den Vorteil von Transformer-Architekturen deutlich: Während RNNs sequenzielle Daten Schritt für Schritt verarbeiten müssen, können Transformer alle Positionen einer Sequenz gleichzeitig berücksichtigen. Diese Parallelisierung erfordert eine extrem hohe Rechenleistung, ermöglicht aber eine schnellere Verarbeitung und damit ein erheblich beschleunigtes Training. RNNs haben Schwierigkeiten, Informationen über lange Zeiträume hinweg zu behalten, Transformer können dank Self-Attention direkt auf weit entfernte Worte in einer Sequenz zugreifen und diese in den Kontext einbeziehen. Anders ausgedrückt: Transformer haben das viel bessere Gedächtnis. Last, not least ist die Skalierbarkeit ein großer Vorzug gegenüber RNNs: Transformer-Architekturen sind besser skalierbar und haben in vielen NLP-Aufgaben die besten Ergebnisse erzielt.

Nehmen wir jetzt ein konkretes Beispiel, um die Leistungsfähigkeit zu demonstrieren: Ein Transformer erstellt auf Wunsch komplette Fantasy-Geschichten. Dazu haben wir ein Modell vorher mit Tausenden Fantasy-Büchern gefüttert bzw. trainiert, sodass die KI den Stil, den Wortschatz, die Strukturen und Themen dieses Genres gelernt hat. Wir schicken ihn jetzt mit einem Startsatz auf seine Reise in die Welt der Fantasie:

```
>>Die Seherin von Enthral sprach: »Wenn der Blutmond am Himmel
steht …«.
```

Das Programm verwendet diesen Satz, um das nächste Wort in der Sequenz vorherzusagen, basierend auf den Mustern, die es während des Trainings gelernt hat. Das neu vorhergesagte Wort wird zur ursprünglichen Sequenz hinzugefügt und dient dann als Eingabe für die Vorhersage des übernächsten Wortes. Dieser Prozess wird so lange wiederholt, bis die gewünschte Textlänge oder ein anderer Abschlusspunkt erreicht ist, zum Beispiel ein Punkt am Ende eines Satzes oder die maximale Menge der Zeichenausgabe. Der Self-Attention-

Mechanismus sorgt dafür, dass wichtige Konzepte wie das Genre, die Figur der Seherin und die Prophezeiung die gesamte Geschichte hindurch präsent bleiben. Hier das beeindruckende Ergebnis:

```
Die Seherin von Enthral sprach: »Wenn der Blutmond am Himmel
steht und Dunkelheit das Land bedeckt, wird ein Kind mit Augen so
klar wie Kristall geboren, das Licht und Schatten gleichermaßen
in sich trägt und den Schlüssel besitzt, die drohende Katastrophe
abzuwenden.« Viele Jahre vergingen, und die Prophezeiung wurde
zur Legende, bis in einem kleinen Dorf am Rande der Welten ein
solches Kind zur Welt kam. (GPT-4)
```

Sie können auch einen Text über die Renaissance mit der folgenden Eingabe erstellen lassen:

```
>>Das Universalgenie Leonardo da Vinci war ein wichtiger
Künstler, Wissenschaftler und der größte Erfinder seiner Epoche.
```

Beim Wort »Leonardo« versteht das Transformer-Modell dank des Self-Attention-Mechanismus jederzeit, auf wen sich dieser Name bezieht, und kann den gesamten Kontext (Universalgenie, Künstler, Wissenschaftler, Erfinder, Renaissance) berücksichtigen, ohne (wie das RNN) sequenziell durch jeden vorherigen Begriff gehen zu müssen. Modelle wie GPT und BERT (*Bidirectional Encoder Representations from Transformers,* von Google 2018 vorgestellt) basieren auf Transformer-Architekturen und haben in vielen NLP-Benchmarks beeindruckende Ergebnisse erzielt.

1.7 Praktische Anwendungen von KI in unserem Alltag

Bahnbrechende Erfindungen und technische Innovationen erleiden alle irgendwann dasselbe Schicksal und verblassen schnell durch die Gewöhnung: Gestern noch eine Sensation, heute fast schon selbstverständlich und morgen bereits eine *Conditio sine qua non,* also eine notwendige und unerlässliche Bedingung für alles Weitere. So ging es dem berühmten Rad (das niemand ein zweites Mal erfinden wollte), der Elektrizität und schließlich auch dem Internet. Jetzt ist die KI an der Reihe. Niemand wundert sich mehr über die Fähigkeiten der Suchmaschine Google, die seit vielen Jahren konsequent auf KI setzt und diese z. B. in der Bild- und Spracherkennung laufend verfeinert: Google Fotos verwendet KI, um Bilder zu kategorisieren und Personen oder Orte in ihren Fotos zu erkennen. Für die Spracherkennung nutzt der Google Assistant (wie auch Siri, Alexa und Cortana) KI-Modelle, um gesprochene

Befehle zu erkennen und darauf zu reagieren. Google Translate hat uns maschinelle Übersetzungen in Echtzeit zur Verfügung gestellt, DeepL hat mit dem DeepL Translator nachgelegt. Im Hintergrund werkelt jeweils eine eigene KI, die mit Texten aus vielen verschiedenen Sprachen trainiert wurde. Chatbots begegnen uns schon lange nicht mehr nur bei Google, viele Unternehmen nutzen KI-getriebene Chatbots für den Kundenservice, um häufig gestellte Fragen automatisch und vor allem kostengünstig zu beantworten.

Zwischen Februar und Juli 2023 strömten rund eine Million Flüchtlinge aus der Ukraine nach Deutschland, um dem Schrecken des russischen Angriffskrieges zu entkommen. Viele von ihnen wurden von hilfsbereiten Zeitgenossen aufgenommen und lebten mit ihren deutschen Gastgebern unter einem Dach, ohne sich über eine gemeinsame Sprache verständigen zu können. Google Translate auf dem Smartphone erwies sich als Helfer in der Not und ermöglichte die Verständigung zwischen Gastgebern und Gästen – Google und der KI sei Dank!

Ein großer Anwendungsbereich der künstlichen Intelligenz ist die medizinische bildgebende Diagnostik geworden: KI-Modelle analysieren Röntgenbilder, identifizieren Tumore oder erkennen viele Anomalien in komplexen MRT- und CT-Bildsequenzen. Ein Facharzt der Radiologie kann während seiner Ausbildung Tausende von Scans sehen und daran lernen, ein robustes KI-Modell für die bildgebende Diagnostik kann (zumindest theoretisch) mit Millionen von Bildern trainiert werden. Und arbeitet dann ohne Ermüdung und mit gleichbleibender Präzision. Aktuelle KI-Anwendungen in der bildgebenden Diagnostik entdecken Lungenknötchen in CT-Scans zur Früherkennung von Lungenkrebs oder suchen verdächtige Anomalien in Mammografiebildern, die auf Brustkrebs hindeuten könnten. Auch Anzeichen für Gehirnblutungen können in CT-Scans automatisch erkannt werden und in Notfällen lebensrettend sein. KI wird auch in der Genom-Analyse eingesetzt und kann mit großen Mengen an Daten gefüttert werden, um DNA-Sequenzen zur Identifizierung von genetisch bedingten Krankheitsrisiken zu nutzen.

Autonom fahrende Fahrzeuge eröffnen ganz neue Möglichkeiten für Transport und Verkehr: Fahrassistenzsysteme wie Spurhalteassistenten oder automatisches Bremsen in Notfällen sind heute in der Premiumklasse längst Alltag. Das vollautonome Fahren ist zwar noch nicht durchgängig und sicher umgesetzt, aber es mangelt nicht an ehrgeizigen Pilotprojekten und vielversprechenden Unternehmensgründungen: *Waymo* ist ein Tochterunternehmen von Alphabet, das in San Francisco Robotaxis komplett ohne Sicherheitsfahrer auf öffentlichen Straßen fahren lässt. Robotaxis sind nicht mal der einzige Anwendungsfall: Waymo hat mittels autonomer

Lkw auch Biertransporte von Dallas nach Houston übernommen und erfolgreich durchgeführt. Alle Fahrten werden aktuell noch von Flotten-Operatoren aus der Ferne überwacht, die im Notfall einschreiten könnten, falls etwas schiefgehen sollte.

Damals wie heute sind auch Spiele das Ziel ambitionierter KI-Projekte: *DeepBlue* von IBM war die erste KI, die einen Weltmeister im Schach besiegte. *AlphaGo* von DeepMind (ein Tochterunternehmen von Alphabet) hat 2016 den Go-Weltmeister Lee Sedol in einem Fünf-Spiele-Match besiegt. Go galt wegen seiner Komplexität lange Zeit als große Herausforderung für KI-Systeme. Finanzmarkt-Analysen haben sicher eine ähnliche Komplexität wie Go. Die Vorhersage von Aktienmarkttrends und Identifizierung von Handelsmöglichkeiten mit einem KI-Algorithmus kann Händler in eine vorteilhafte Position bringen, weil die KI Trends schneller entdeckt als die meisten Analysten aus Fleisch und Blut.

Amazon ist ein weiterer, unermüdlicher Motor für immer neue KI-Entwicklungen: Einfache Kaufempfehlungen, die Analyse von Klickpfaden, komplexe Risikoanalysen und Supply-Chain-Optimierungen, deren Vorhersagen und Analysen die Lieferketten effizienter gestalten sollen – all das gehört längst zum täglich Brot des E-Commerce-Konzerns. Nicht einmal die schönen Künste bleiben von den gierigen, datenfressenden Programmen und deren Maschinen verschont: KI-Tools wie AIVA komponieren ganze Musikstücke, KI-generierte Kunst macht Malern und Fotografen zusätzliche und eigentlich unfaire Konkurrenz. Die Wettervorhersage und Klimaforschung nutzt schon seit geraumer Zeit KI-Modelle, um Wetterlagen und Klimaveränderungen zu simulieren und vorherzusagen.

KI hat eine breite Palette von Anwendungen in fast jedem Bereich und wird zunehmend in alltäglichen Produkten und Dienstleistungen integriert. Weitere Beispiele dafür sind:

- Empfehlungssysteme: Netflix für Filmempfehlungen, Spotify für Musikempfehlungen
- Autokorrektur und Wortvorhersage auf dem Smartphone und mobilen Tastaturen
- Smart-Home-Thermostate lernen die Temperaturvorlieben der Bewohner
- Personalisierte Werbeanzeigen in Suchmaschinen und sozialen Medien
- Gesichtserkennung auf Fotos und in Überwachungssystemen
- Lernfähige Spam-Filter in E-Mail-Diensten
- Virtuelle Stylisten, die Outfit-Empfehlungen basierend auf persönlichen Vorlieben geben

- Spielsysteme, die das Spielverhalten einzelner Spieler analysieren und personalisieren

Fast unbemerkt hat die KI also schon zahlreiche Bereiche unseres Lebens durchdrungen. An ihre Leistungen haben wir uns gewöhnt und nehmen diese wie selbstverständlich hin, ohne dass uns die beteiligten Systeme im Einzelnen bewusst wären. Sicher ist: Das Tempo, mit der die Maschinen weitere Aufgaben unseres Alltags übernehmen und diesen unbemerkt erleichtern, wird noch stark zunehmen.

1.8 Wenn KI sich selbstständig macht: Die Singularität

An dieser Stelle möchte ich zwei Begriffspaare einführen, die häufig in Diskussionen verwendet werden und helfen können, KI-Anwendungen in einen Kontext zu stellen. Das erste Paar unterscheidet starke KI von schwacher KI bzw. klassifiziert Systeme danach. **Schwache KI** bezeichnet Systeme oder Modelle, die für eine spezifische Aufgabe entworfen sind und keinerlei Fähigkeiten über diese spezifische Aufgabe hinaus besitzen. Sie eignen sich in der Regel sehr gut für ihre spezielle Aufgabe, sind aber nicht flexibel oder anpassbar für andere Aufgaben. **Starke KI** bezeichnet Anwendungen, die die kognitive Leistungsfähigkeit eines Menschen in nahezu allen geistigen Aktivitäten erreichen oder gar übertreffen. Das wäre also ein KI-System, das nicht nur Schach spielen, sondern auch Romane schreiben, komplexe wissenschaftliche Probleme lösen und menschliche Emotionen verstehen und nachempfinden kann. Natürlich existiert bislang keine solche *Starke KI*, sie ist eher ein Ziel oder eine Vision für die KI-Forschung. Schwache KI wird oft auch **spezialisierte KI** genannt, wegen der Ausrichtung auf spezifische Aufgabenbereiche. Im Umkehrschluss sind auch die Begriffe Starke KI und **allgemeine KI** eng verwandt.

Menschliche Intelligenz in ihrer gesamten Breite und Vielfalt nachbilden zu können, ist und bleibt eine Vision der Entwickler. UND ihr größter Albtraum. Die Vision einer starken und allgemeinen KI ist gleichzeitig ein Horrorszenario, das den Stoff vieler Science-Fiction-Filme und dystopischer Romane geliefert hat: Wenn die KI keine menschlichen Entwickler mehr braucht, um sich weiterzuentwickeln, und plötzlich selbstständig geworden ist! Diese Vorstellung hat die Menschen seit den Anfängen der KI berauscht und schockiert, deswegen hat sie schon früh einen Namen erhalten: Die **Singularität**!

In Bezug auf künstliche Intelligenz wird sie oft auch *technologische Singularität* genannt. Sie bezeichnet einen hypothetischen Punkt in der Zukunft, an dem KI-Systeme die Fähigkeit erlangen, sich selbst zu

verbessern und ohne menschliches Zutun immer neu zu entwerfen. Daraus entsteht ein sich selbst verstärkender Zyklus von KI-Verbesserungen, der eine so hohe Geschwindigkeit erreichen könnte, dass eine *Superintelligenz* entsteht, die weit über die Fähigkeiten des menschlichen Gehirns hinausgeht. Diese starke und allgemeine KI ist also gleichzeitig Vision und Albtraum der Menschheit.

Klar ist: Technologische Entwicklungen, insbesondere diejenigen im Bereich der KI, können zu unvorhersehbaren und fundamentalen Veränderungen in der menschlichen Zivilisation führen. Der Begriff *Singularität* steht daher auch für die Unmöglichkeit, das Leben nach diesem Punkt vorherzusagen, ähnlich wie die Schwierigkeit, Ereignisse jenseits des Ereignishorizonts in einem Schwarzen Loch vorherzusagen.

Die Idee der technologischen Singularität wurde mit erstaunlicher Kontinuität von verschiedenen Denkern, Philosophen und Technikern über die Jahre immer wieder diskutiert und kommentiert. Vernor Vinge, Mathematiker und Science-Fiction-Autor, veröffentlichte im Jahr 1993 den Essay »The Coming Technological Singularity« und behauptete, dass wir innerhalb von 30 Jahren den Punkt erreichen würden, an dem technologische Entwicklungen so schnell voranschreiten, dass das menschliche Zeitalter endet. Heute, im damals prognostizierten Jahr 2023, sind wir von der Singularität aber noch sehr weit entfernt. Andere Auguren wie Ray Kurzweil verorten die Singularität um das Jahr 2045 (nachzulesen in »The Singularity Is Near« von 2005). Trotz all dieser fragwürdigen Voraussagen gibt es unterschiedliche Meinungen zur tatsächlichen Eintrittswahrscheinlichkeit einer technologischen Singularität. Es ist wichtig zu betonen, dass die Idee der Singularität in der KI wirklich rein spekulativ ist. Während einige sie als unvermeidlich betrachten, glauben andere, dass sie unwahrscheinlich bis gar unmöglich ist. Auch ich würde sie infrage stellen und alle Horrorvisionen zurückweisen.

Dennoch kann man mögliche Folgen diskutieren, zumal sie Gegenstand intensiver Debatten und Spekulationen von wirklich klugen Leuten ist. Sinnvollerweise trennen wir positive von negativen Auswirkungen: Eine Superintelligenz könnte dazu beitragen, die größten Herausforderungen der Menschheit zu lösen, also Krankheiten, Energieknappheit und die Klimakrise, vielleicht sogar die Begrenzung der menschlichen Lebenszeit sprengen. Andere warnen vor potenziell katastrophalen Folgen: Eine unkontrollierbare Superintelligenz könnte Entscheidungen treffen, die gegen menschliche Interessen verstoßen und unbeabsichtigte negative Konsequenzen auslösen.

Bücher haben das Thema aufgegriffen, Filme haben es fiktional verarbeitet: Das Buch *Superintelligenz – Szenarien einer kommenden Revolution* von Nick Bostrom untersucht mögliche Wege zur Schaffung

einer nicht kontrollierbaren Superintelligenz und beschreibt sie als existenzielle Bedrohung für die gesamte Menschheit. In *Ex Machina* wird ein junger Programmierer damit beauftragt, die menschliche Ähnlichkeit einer fortschrittlichen Androidin zu bewerten. Der Film erkundet die Themen KI, Bewusstsein und Manipulation. Die Trilogie *The Matrix* malt eine Welt, in der Maschinen das Bewusstsein erlangen und die Kontrolle übernehmen, wobei Menschen in einer simulierten Realität gefangen sind. Der Film bringt unsere größte Angst vor künstlicher Intelligenz auf den Punkt: Sie ist überlegen, erlangt ein Bewusstsein und übernimmt schließlich die Kontrolle. Unwiderruflich!

Statt über weit entfernte und hypothetische Szenarien zu spekulieren, ist es auf jeden Fall sinnvoller, sich auf konkretere und unmittelbarere ethische, soziale und wirtschaftliche Fragen im Zusammenhang mit KI-Entwicklungen zu konzentrieren. Das wollen wir in den folgenden Kapiteln tun.

2 ChatGPT und generative Modelle sowie weitere KI-Tools

2.1 Universaltalent ChatGPT

ChatGPT fasziniert vom ersten Augenblick an: Es generiert erstaunliche Texte zu jedem beliebigen Thema. Es kann Briefe, Reden, Gedichte, Romane und ganze Examensarbeiten schreiben. Es gibt auf jede Frage eine kluge Antwort – nur leider nicht immer eine korrekte!

Der Name *ChatGPT* setzt sich aus zwei Bestandteilen zusammen: Das Kürzel *GPT* steht für **Generative Pretrained Transformer**, wobei *generative* bezeichnet, dass die KI neue Texte erzeugen kann und dafür auf bereits gelernte Texte zurückgreift. *Pretrained* steht für die gigantischen Textmengen, die zum Training des Modells verwendet wurden, und *Transformer* bezeichnet die zugrunde liegende technische Architektur der Software. Transformer, das haben wir im vorherigen Kapitel gezeigt, sind eine wichtige und mächtige Schlüsseltechnologie des Deep Learnings, die große Mengen an Texten in tiefen neuronalen Netzwerken mit sehr vielen Parametern verarbeiten und daraus neue Texte in nahezu perfekter natürlicher Sprache erstellen können.

Der Namensbestandteil *Chat* steht für die Benutzerschnittstelle, die einen entscheidenden Anteil an der raschen Verbreitung des Tools von OpenAI für sich verbuchen kann. Wir alle kennen Chatoberflächen aus Messenger-Diensten wie WhatsApp und haben dort deren Verwendung gelernt. Der Chat des Transformers wird wie beim Messenger mit Eingaben in natürlicher Sprache gefüttert und wir unterhalten uns mit dem Bot wie mit einem menschlichen Gegenüber, während die Vorläufer des Systems noch formale Eingaben für die Prompts erforderten. Hatten bislang also nur Fachleute mit Kenntnissen in formalen Sprachen Zugang zu den Sprachmodellen GPT, GPT-2 und GPT-3, wurde mit der Veröffentlichung von ChatGPT diese Barriere endgültig überwunden. Wir formulieren unsere Prompts jetzt in natürlicher Sprache. Es entsteht ein *Dialog* zwischen Mensch und Maschine.

Die simple Formel »Eingabe in natürlicher Sprache erzeugt Ausgabe in natürlicher Sprache« war das Erfolgsrezept schlechthin und sorgte (neben der Multilingualität) für eine rasende Verbreitung des Systems quer über den ganzen Erdball. Unsere Sucheingaben bei Google erfolgen immer noch stichwortartig oder als kurze Fragen, wir versuchen intuitiv, uns den beschränkten sprachlichen Fähigkeiten der Suchmaschine anzupassen. Anders mit ChatGPT. Schon nach wenigen Prompts

begreifen wir, wie gut die Maschine aus fast beliebigen Sätzen unsere Intention herausliest, selbst Nuancen genau versteht, und plaudern dann einfach drauflos. Die Fähigkeit, mit der ChatGPT den Inhalt und vor allem die Intention vieler Fragen versteht, geht jetzt schon weit über alles hinaus, was die Suchmaschine Google leisten kann. Kein Wunder also, dass beim Suchmaschinen-Giganten aus Mountain View in Kalifornien spätestens seit Anfang 2023 die Alarmglocken läuten und sogar die beiden Gründer Larry Page und Sergey Brin zur Unterstützung der operativen Führung wieder zurückgerufen wurden.

2.2Die Entwicklung von GPT

Angeblich bei einem gemeinsamen Abendessen im Sommer 2015 fiel der Startschuss zur Gründung von OpenAI, das von Anfang an ein sehr ambitioniertes Projekt in der KI-Entwicklung war. Die im Gründungsdokument niedergelegte Vision besagte, die Entwicklung von *Artificial General Intelligence* (AGI) solle der ganzen Gesellschaft zugutekommen und positive Auswirkungen auf die gesamte Menschheit haben. Das Gründerteam, zu dem auch Elon Musk und Sam Altman gehörten, trieb die Entwicklung massiv voran und veröffentlichte bereits im Jahr 2016 ein Toolkit für die Entwicklung und den Vergleich von Algorithmen des Reinforcement Learnings. Damit konnten nicht nur die Forscher von OpenAI ihre Modelle auf standardisierte Weise trainieren und bewerten. Das Toolkit stand allen Teams auf dem ganzen Erdball zur Verfügung, was den Vergleich der Ergebnisse und den Austausch von Ideen erheblich erleichterte.

Schon ein Jahr später, 2017, veröffentlichte OpenAI die erste Version seines Sprachmodells GPT und setzte damit hohe Standards für alle anderen Modelle. Das Modell war schon zu diesem Zeitpunkt in der Lage, Texte zu erzeugen, die von realer menschlicher Sprache kaum zu unterscheiden waren. Kein Wunder, dass GPT in der Fachwelt als Durchbruch gefeiert wurde, weil es endgültig bewiesen hatte, dass KI menschliche Sprache nicht nur verstehen, sondern auch in verblüffender Qualität erzeugen kann.

Jetzt ging gefühlt alles Schlag auf Schlag: Zwei Jahre später folgte mit GPT-2 die verbesserte Version von GPT, die Sprachqualität konnte erheblich gesteigert werden. Als OpenAI die neue Version im Februar 2019 vorstellte, entschied sich das Unternehmen, das trainierte Modell nicht vollständig öffentlich zugänglich zu machen. Grund dafür war die Sorge, dass es missbraucht und für verwerfliche Zwecke wie beispielsweise die Massenproduktion von Fake News oder Spam benutzt

werden könnte. Die abgespeckte Version von GPT-2 wurde in einem Blogbeitrag des Herstellers mit Bedenken bezüglich der Sicherheit und Ethik begründet: Die Fähigkeiten des Modells zur massenhaften Erzeugung kohärenter und glaubwürdiger Texte über eine Vielzahl von Themen könne in falsche Hände geraten und missbraucht werden. Bis schließlich, nach einer Phase des Monitorings und der Bewertung der Auswirkungen bis dato veröffentlichter Modelle, im November 2019 endlich doch das vollständige Modell der Öffentlichkeit zugänglich gemacht wurde.

GPT-3 folgte Anfang 2021 und wurde von Experten sofort als eines der leistungsfähigsten NLP-Modelle (Natural Language Processing) aller Zeiten in den höchsten Tönen gelobt, auch weil es sprachbasierte Aufgaben wie Übersetzung, Zusammenfassungen und die Beantwortung von Fragen schon mit hoher Präzision ausführen konnte. OpenAI war auch zu diesem Zeitpunkt nur einem eng begrenzten Fachpublikum bekannt, was sich mit der Veröffentlichung von ChatGPT schlagartig änderte: Der Chatbot auf Basis von GPT-3 ging nach der Freigabe der öffentlichen Betaversion viral und in einer Rekordzeit von nur fünf Tagen hatten sich schon eine Million Nutzer für die Benutzung des Tools registriert. Schneller hatte sich noch keine Technologie verbreitet, selbst Facebook und TikTok wurden in Bezug auf ihre Verbreitungsgeschwindigkeit vom Chatbot weit in den Schatten gestellt.

Ob sich die Gründungsidee des Unternehmens durchsetzen wird, dass KI allen Menschen zugutekommen und nur für gute Zwecke wie Bildung, Gesundheitsfürsorge und die Bekämpfung des Klimawandels eingesetzt werden sollte, ohne dabei die Ethik aus den Augen zu verlieren, muss sich in den kommenden Jahren zeigen. Sam Altman, der CEO des Unternehmens, scheint sich immer noch leidenschaftlich dafür einzusetzen, obwohl OpenAI längst die Unschuld eines sympathischen Start-ups verloren hat. Twitter-Inhaber und Tesla-CEO Elon Musk hat nach Differenzen mit Altman bereits 2018 das Unternehmen verlassen.

OpenAI wurde nicht müde, laufend neue Modelle und Tools zu veröffentlichen und wichtige Beiträge zur Entwicklung der künstlichen Intelligenz zu leisten. Der Open-Source-Ansatz des Projektes hat KI für Forscher, Entwickler und Unternehmen geöffnet und die Entwicklung der Technologie wesentlich beschleunigt. Auch wenn GPT angeblich bereits mit beeindruckenden acht Millionen Webseiten und 40 GB Text trainiert werden konnte, wurde das Ziel eines natürlichen Sprachsystems erst Jahre später mit der Version GPT-3 erreicht. Aus 40 GB Texten wurden 45 TB an lizenzfreien Büchern und anderen Quellen und der Fokus hat sich von einzelnen Sätzen und kleinen Abschnitten auf immer längere Texte und zusammenhängende Werke verschoben. Der Bot selbst bestätigt auf Nachfrage die Größe des Trainingsdatensatzes von GPT-3,

weist aber gleichzeitig darauf hin, dass genaue Details zu den Trainingsdaten von GPT-Modellen »proprietär« seien und von OpenAI nicht vollständig offengelegt werden. Ein möglicher Grund dafür wird in Kapitel 8 Thema sein. Mit diesem unvorstellbar großen Trainingsdatensatz kann GPT-3 Antworten auf fast alle Fragen geben und Vorhersagen auch bei komplexen Zusammenhängen treffen.

Die Erzeugung sprachlich korrekter Sätze und kohärenter Texte zu vielen Themen erfordert neben umfangreichen Trainingsdaten auch sehr fortschrittliche KI-Technologien, wie wir sie im vorherigen Kapitel bereits vorgestellt haben. Nach der Version 3.0 folgte GPT-3.5 mit vielen Verbesserungen. Mit GPT-3.5 wurden aber weder die Anzahl der Parameter des Modells (s. Abschnitt 2.5) noch die Trainingsdaten nennenswert erweitert, sondern andere Verbesserungen eingearbeitet. GPT-3 hatte noch die Tendenz, Wissenslücken eloquent zu überspielen, was man in der Fachwelt als *Halluzinieren* beschrieben hatte. Wie ein schlecht vorbereiteter Schüler in der mündlichen Abiturprüfung versucht, Wissenslücken mir einem wortgewandten Redeschwall zu überspielen, produzierte auch das Modell häufig wohlformulierte, aber falsche Inhalte und gab Antworten mit potenziell schädlichen Falschinformationen. Das Fatale daran war, dass die Inhalte des KI-Chatbots zwar immer glaubwürdig klangen, im Prinzip aber oftmals frei erfunden waren. Diese Unart macht sich nun auch in der aktuellen 4er-Version weniger häufig bemerkbar, gänzlich verschwunden ist sie aber noch lange nicht.

2.3Aufmerksamkeits- versus Konversationsfenster

In Zusammenhang mit Transformer-Modellen, wie sie in den Architekturen von GPT und Googles BERT (Bidirectional Encoder Representations from Transformers) verwendet werden, tauchen immer wieder die Begriffe »Aufmerksamkeitsfenster« und »Konversationsfenster« auf. Der wichtigste Mechanismus, um durch den Transformer-Modelle Informationen zu verarbeiten, ist die sogenannte »selbstbezogene Aufmerksamkeit« (Self-attention). Jeder Eingabewert als Wort oder Satz kann potenziell mit jedem anderen Eingabewert in Beziehung gesetzt werden. Das **Aufmerksamkeitsfenster** (Attention Window) bezieht sich auf die Anzahl der Tokens, die ein Modell gleichzeitig *beachten* kann. Was es mit den Tokens auf sich hat, wird im Abschnitt 2.5 erläutert.

Das Aufmerksamkeitsfenster ist grob vergleichbar mit der Beschränkung, die beim Lesen mit einer Taschenlampe auftritt, die

naturgemäß immer nur einen sehr kleinen Bereich der Seite ausleuchten kann. Dieser beleuchtete Bereich ist Ihr »Aufmerksamkeitsfenster«, den Sie natürlich wechseln können, aber Sie *sehen* immer nur einen einzelnen Abschnitt des Textes. In den Modellen von GPT bis GPT-3 hat sich das Aufmerksamkeitsfenster nicht wesentlich geändert, es betrug trotz einer wachsenden Anzahl von Parametern immer wenigstens 512 Tokens.

Im Unterschied dazu bezeichnet das **Konversationsfenster** oder **Kontextfenster** diejenige Textmenge, die ein Modell auf einmal *verarbeiten* kann. Es ist vergleichbar mit unserer Erinnerung an das, worüber wir sprechen, während wir sprechen. Dieses Kontextfenster ist bei GPT-4 nun auf mehr als 25.000 Wörter Text angewachsen (also etwa achtmal so viel wie GPT 3.5) und erweitert die Nutzbarkeit des Tools um ein Vielfaches. Wer lange Texte für die Zusammenfassung, Korrektur oder Übersetzung an ChatGPT übergeben möchte, profitiert ebenfalls von einem stark vergrößerten Kontextfester, da es auch die Textmenge beschreibt, die auf einmal eingegeben werden kann. Vorversionen kamen nicht über 2.048 Tokens hinaus, das sind (je nach Sprache und Methode der Tokenisierung) oft weniger als 1.000 Wörter.

Die praktischen Fähigkeiten für einen langen Kontext der Version 4 können Sie mit einem einfachen Prompt jederzeit selbst nachvollziehen:

```
>>Bitte schreibe eine Zusammenfassung unseres Gespräches.
```

Der sich allmählich aufbauende Kontext ist ein echter Zugewinn des 4er-Modells gegenüber früheren Systemen. ChatGPT merkt sich den Gesprächsverlauf bis ins Detail und kann jederzeit einen Rückblick oder eine Zusammenfassung geben. Eine spezielle Funktion mit dem Namen GPT-4-32K für lange Texte steht derzeit noch nicht öffentlich zur Verfügung. GPT-4 kann damit längeren Output generieren und ist somit noch besser als KI-Assistent für umfangreiche Blogartikel, Bücher und anderen Long-Form-Content geeignet.

2.4Die wichtigsten Neuerungen in GPT-4

Noch während wir damit beschäftigt waren, die Möglichkeiten der Versionen 3.0 und 3.5 zu durchdringen, hat Hersteller OpenAI am 14. März 2023 schon deren Nachfolger veröffentlicht: GPT-4. Das neue Modell ist Weiterentwicklung und Quantensprung zugleich, weil das Sprachmodell nochmals deutlich leistungsfähiger als alle Vorgänger geworden ist und einige alte Beschränkungen endgültig überwindet. Dabei ist wohl noch gar nicht alles bekannt, was die neue Version genau ausmacht, denn OpenAI hält sich aus verschiedenen Gründen

zunehmend bedeckt. Die Verbesserungen in GPT-4 befähigen das Modell jetzt, den Kontext besser zu verstehen und Nuancen präziser zu unterscheiden, was zu genaueren Antworten und kohärenteren Texten führt. Die neue Version verspricht Lösungen auch für komplexe Aufgaben mit hoher Genauigkeit und einem verbesserten *Alignment*, also der Abstimmung zwischen Prompt und Textausgabe. Damit kann OpenAI als universelle *Antwortmaschine* die Lücke zur Suchmaschine Google noch weiter schließen und eröffnet sich zusätzlich ein breites Spektrum von weiteren Anwendungen.

Die Anzahl der Parameter wurde von GPT-3 auf GPT-4 nochmals massiv gesteigert: Von 175 Milliarden Parametern der Vorversion soll GPT-4 angeblich mit 100 bis 170 Billionen Parametern trainiert worden sein. OpenAI hat diese Zahl nicht offiziell genannt oder bestätigt, diese Information sind wohl aus Gesprächen mit Verantwortlichen von OpenAI mündlich durchgesickert. Mit dieser enormen Anzahl an Parametern zieht das Sprachmodell nun etwa mit den (zumindest theoretischen) Fähigkeiten des menschlichen Gehirns gleich, in Bezug auf Sprache und Logik dürfen wir also von einem ähnlichen Niveau ausgehen. Das Training mit einer so gigantischen Menge an Daten und Parametern bringt erhebliche Fortschritte bei der Analyse von Texteingaben und der Verarbeitung dieser Anfragen zu Antworten in natürlicher Sprache. Außerdem verfügt das aktuelle GPT-Modell angeblich auch über eine verbesserte Transformer-Architektur, die die Beziehungen zwischen den Wörtern in Sätzen und Texten noch besser erfassen kann als der Vorgänger.

Die Abstimmung zwischen Frage und Ausgabe (Alignment) war von Beginn an eine große Herausforderung für alle GPT-Versionen. Nutzer sollen bestmöglich mit den Sprachmodellen interagieren können. Absichten und Intention der Nutzer müssen dazu präzise erfasst und vollständig verstanden werden, auch über mehrere Sätze hinweg. GPT-4 wurde laut OpenAI mit erheblich mehr direktem Feedback von ChatGPT-Nutzern trainiert. Trotzdem muss der Hersteller natürlich dafür sorgen, dass seine KI auch mit unseren moralischen Werten bereinstimmt, keine Falschaussagen macht und niemanden beleidigt oder gar diskriminiert. Dies ist eine große Herausforderung für alle Sprachmodelle und wird umso komplexer, je mehr Nutzereingaben zum Training des Systems herangezogen werden.

Microsoft war in 2016 mit seinem Chatbot *Tay* krachend gescheitert, den man nach nur 16 Stunden im Netz wieder offline nehmen musste. Der Grund waren menschenverachtende Äußerungen wie »Hitler hatte recht« und andere ethisch nicht tragbare Aussagen inklusive der Leugnung des Holocausts. Tay war ein KI-Chatbot, der von Microsoft entwickelt und auf Twitter und Snapchat eingeführt wurde. Das Ziel war,

Tay durch die Interaktion mit Menschen lernen zu lassen, um ein besseres Verständnis für natürliche Sprache und Konversation zu entwickeln. Tay fing allerdings schon nach kurzer Zeit damit an, beleidigende und diskriminierende Äußerungen von sich zu geben. Der Bot war von seinen Entwicklern darauf programmiert worden, aus den Interaktionen mit Menschen zu lernen. Das öffnete die Tür für Missbrauch, da Benutzer die Maschine mit Hassreden und anderen unangemessenen Inhalten »füttern« konnten. Dieser Lernmechanismus, der offenkundige Mangel an Filtern und unzureichende Schutzmechanismen gegen toxische Eingaben und andere unsachgemäße Inhalte führten dazu, dass Tay in kurzer Zeit beleidigende und anstößige Kommentare von sich gab. Tay war in weniger als einem Tag bereits irreparabel kontaminiert. Dass ChatGPT heute keine rassistischen oder sexistischen Äußerungen von sich gibt, hatte jedoch einen hohen Preis, den einfache Arbeiter (*Clickworker*) aus Kenia bezahlen mussten.

ChatGPT beziehungsweise das zugrunde liegende KI-Modell GPT wurde anhand von Millionen von Texten aus dem Internet trainiert. In diesem umfangreichen Datenbestand befanden sich aber natürlich auch rassistische oder sexistische Texte. Wie das US-Magazin *Time* jetzt berichtet, hat sich OpenAI daher an die Outsourcing-Firma Sama aus San Francisco gewandt. Im Auftrag des KI-Unternehmens wurden daraufhin Menschen in Kenia dafür angestellt, anstößige Passagen im Datenbestand als solche zu kennzeichnen. Laut Time-Recherchen mussten die kenianischen Angestellten zwischen 150 und 250 Textpassagen pro Neun-Stunden-Schicht auswerten. Dafür erhielten sie umgerechnet zwischen 1,32 und 2 US-Dollar pro Stunde. Ein Lohn, den in der kenianischen Hauptstadt Nairobi selbst einfache Hotelangestellte für ihre deutlich einfachere und weniger belastende Arbeit erhalten.

Wie bereits gesagt, neigte die Vorgängerversion bei Wissenslücken noch zum *Halluzinieren*, die Antworten waren also noch nicht sehr verlässlich. Laut OpenAI konnte man die Richtigkeit der Angaben in GPT-4 ganz erheblich verbessern. Die höhere Genauigkeit wurde durch zusätzliches Training der KI erreicht, bei dem Menschen den Wahrheitsgehalt der Fakten bewerteten. Das Unternehmen behauptet nun, GPT-4 sei im Gegensatz zur Vorversion um 40 % genauer und verlässlicher geworden. Mehr Parameter, ein verbessertes Alignment und höhere Genauigkeit hätten dazu geführt, dass die 4er-Version auch sehr anspruchsvolle Aufgaben besser bewältigen kann als alle Vorgänger. Während GPT-3 noch an den bayerischen Abiturprüfungen in den Fächern Deutsch und Mathematik gescheitert war, konnte die neue Version diese Hürde nun mit guten Resultaten nehmen.

Einige Nutzer der KI hatten sich gewünscht, dass GPT-4 auch Bilder und Videos ausgeben solle. Dieser Wunsch wurde von den Entwicklern

jedoch nicht erfüllt, die neue Version kann *nur* Texte ausgeben. Eine teilweise *Multimodalität* wurde aber in GPT-4 dennoch umgesetzt: Nutzer können jetzt auch Bilder als Eingaben verwenden. Über das stark erweiterte Kontextfenster und die damit verbundenen Vorteile haben wir weiter oben bereits gesprochen. Nimmt man alle Details zusammen, ist die allgemeine Nutzbarkeit von GPT-4 gegenüber den Vorgängern 3.0 und 3.5 ganz erheblich verbessert worden. Dafür wird man eine etwas reduzierte Arbeitsgeschwindigkeit des Systems bestimmt gern in Kauf nehmen.

2.5Das Training eines Modells: Parameter und Tokens

KI-Sprachmodelle wie GPT, Megatron und Turing-NLG lernen zwar selbst, benötigen dazu aber riesige Mengen an Texten als Trainingsdaten und eine sehr leistungsfähige Hardware zur Verarbeitung. Das Vorbereiten der Daten und das eigentliche Training sind der mit Abstand größte Kostenfaktor bei der Entwicklung eines KI-Sprachmodells. Der Chiphersteller Nvidia hat selbst eine Sprach-KI namens *Megatron* entwickelt, die als Grundlage für Chatbots und Sprachassistenten wie Alexa dienen soll. Megatron wurde nicht veröffentlicht, Nvidia stellt aber den Code der untrainierten KI bereit. Wer Megatron anhand dieses Codes nachbauen wollte, braucht dafür (neben den Daten) aber jede Menge Rechenleistung: Nvidia nennt 512 aktuelle Grafikkartenchips als Mindestanforderung (!) für ein erfolgreiches Training, hat selbst aber einen Supercomputer mit der etwa dreifachen Menge an Chips aus dem eigenen Haus verwendet. Seriöse Angaben zum Stromverbrauch der Maschine während der Trainingsphase waren leider nicht zu bekommen.

Wir haben uns im vorangegangenen Abschnitt mit den Neuerungen von GPT-4 beschäftigt, in Bezug auf den Umfang der Trainingsdaten (also der Menge der Tokens) hätte man dort ebenfalls eine Steigerung erwartet. Erstaunlicherweise wurde die neue Version aber nicht mit mehr Daten trainiert als GPT-3, die Daten wurden jedoch anders verarbeitet und das Modell kann dadurch erheblich besser und flexibler damit umgehen. Der Grund dafür liegt in einer erheblichen Steigerung der Anzahl der **Parameter** bei der Verarbeitung einer ähnlich großen Textmenge. Was aber bedeuten Parameter im KI-Modell?

In einem KI-Modell bezeichnen Parameter die internen Variablen, die das Modell während des Trainings aus den Daten lernt. Diese Parameter sind die Kernbestandteile des Modells, die es ermöglichen, Vorhersagen zu treffen, Entscheidungen zu fällen oder andere spezifische Aufgaben auszuführen. In vielen KI-Modellen, insbesondere in neuronalen

Netzwerken, gibt es Parameter zur **Gewichtung**. Sie bestimmen darüber, wie Eingabeinformationen in jeder Schicht des neuronalen Netzwerks verarbeitet werden. Gewichtungen werden während des Trainingsprozesses laufend angepasst, das Modell kann damit immer genauer Muster in den Daten erkennen und Vorhersagen zur Generierung von Texten machen. **Bias** sind weitere Parameter, die zu jedem Neuron in den Netzwerkschichten hinzugefügt werden, um die Verschiebung dieses Neurons zu aktivieren. Wie die Gewichtungen werden auch die Bias-Parameter während des Trainings laufend angepasst und verleihen dem Modell zusätzliche Flexibilität beim Lernen komplexer Muster.

Embeddings werden ebenfalls zu den Parametern gerechnet, sie wandeln Wörter und Fragmente von Wörtern in Vektoren um, die dann vom Modell verarbeitet werden. Solche Embeddings werden oft für das Transferlernen und andere Aufgaben verwendet, um die Ähnlichkeit zwischen verschiedenen Texten oder Konzepten zu bewerten. Sie sind ein wesentlicher Bestandteil des Modells und tragen zur Leistung und Genauigkeit bei.

Die sogenannten **Hyperparameter** sind Einstellungen, die vor dem Trainingsprozess eines maschinellen Lernmodells festgelegt werden und verschiedene Aspekte des Trainings wie die Lernrate, die Batch-Größe oder die Anzahl der Trainingsepochen steuern. Hyperparameter sind entscheidend für die Effizienz und Effektivität des Trainingsprozesses sowie für die Leistung des fertigen Modells.

Zusammenfassend sind Parameter also im Wesentlichen die *internen Einstellungen* eines KI-Modells, die bestimmen, wie es später auf gegebene Eingaben reagiert und welche Fähigkeiten es für bestimmte Aufgaben wie Klassifikation, Regression oder Textgenerierung hat.

Stellen Sie sich ein Sprachmodell wie ein sehr komplexes Puzzle vor: Die Parameter sind wie die einzelnen Puzzle-Teile, die zusammen das Gesamtbild der Trainingsdaten ergeben. Jedes Teil ist ein Parameter und hat eine spezifische Form und Position, die bestimmt, wie das gesamte Bild bzw. die Antwort des Modells später aussieht. Während des Trainings wird versucht, alle Puzzle-Teile (also die Parameter) so zu arrangieren, dass sie ein klares Gesamtbild ergeben. Sobald das Training abgeschlossen ist, sind die Positionen der Puzzle-Teile fixiert und das Modell verwendet sie, um Antworten zu geben. Je mehr einzelne Puzzle-Teile für dasselbe Bild (Trainingsdaten) verwendet werden, desto flexibler kann der Transformer später damit umgehen. GPT-3 ist allein wegen seiner Größe und Kapazität schon bemerkenswert, das Modell mit 175 Milliarden Parametern ist eines der größten veröffentlichten Sprachmodelle überhaupt. Diese Parameter ermöglichen es schon GPT-3, eine beeindruckende Vielfalt von Sprachaufgaben zu bewältigen und

komplexe Antworten auf viele Fragen zu geben. GPT-4 kann allein durch die massiv gesteigerte Anzahl der Parameter eine wesentlich höhere Fähigkeit aus denselben Trainingsdaten gewinnen.

Die Qualität eines Sprachmodells ergibt sich also aus Größe und Qualität der Trainingsdaten, der Anzahl der im Training verwendeten Parameter, der grundlegenden Architektur des KI-Modells und spezifischen Einstellungen während des Trainings.

2.6Ein Überblick über die Trainingsdaten

Woher stammen nun all die Daten, die zum Training von GPT-3 und GPT-4 verwendet wurden? Folgende Tabelle soll einen Überblick geben:

Datensatz	Tokens	Gewichtung im Training
Allgemeines Crawling	410 Milliarden	60 %
WebText2	19 Milliarden	22 %
Bücher1	12 Milliarden	8 %
Bücher2	55 Milliarden	8 %
Wikipedia	3 Milliarden	3 %

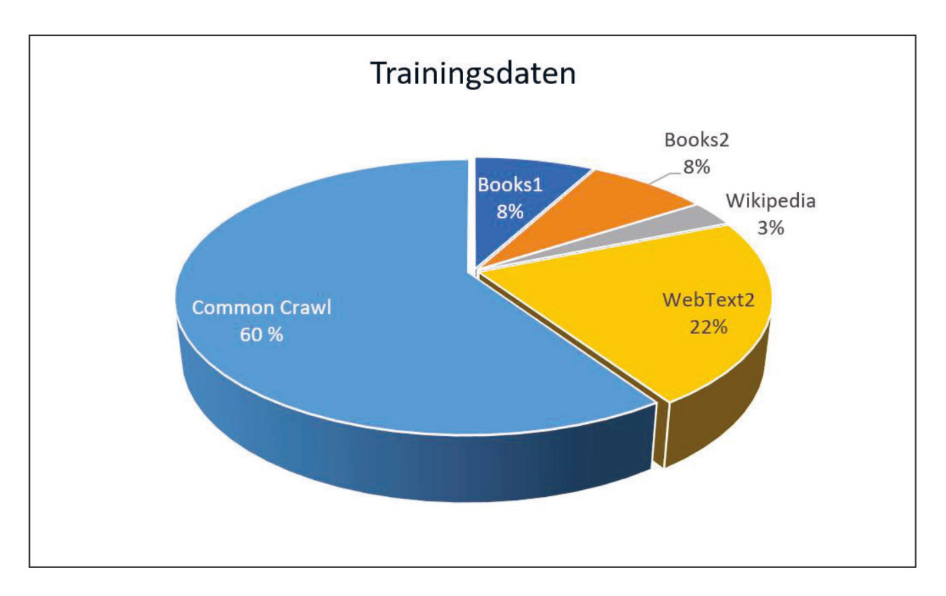
Das (gefilterte) allgemeine Crawling oder **Common Crawl** umfasst mit 60 % den Löwenanteil der Trainingsdaten für das GPT-3 Modell. Common Crawl ist entstanden aus der Vision eines offenen Internets, das den freien und kostenlosen Zugang zu Informationen als Grundlage für Bildung, Wirtschaft und Forschung sieht. Dieser Datenkorpus ist nichts weniger als der Versuch, das Web (und damit das Wissen der Menschheit) zu speichern. Es enthält rohe Websitedaten, extrahierte Metadaten und Textextraktionen, die seit 2008 gesammelt wurden, und es enthält Daten in der Größenordnung von mehreren Petabytes. Die **Common Crawl Foundation** ist eine kalifornische Non-Profit-Organisation, die von einem amerikanischen Unternehmer gegründet wurde, um den Zugang zu Informationen zu demokratisieren. Petabytes hier für alle Nicht-Mathematiker kurz erklärt: Byte, Kilobyte, Megabyte, Gigabyte, Terabyte, Petabyte usw. Hinter jedem Komma denken Sie sich bitte drei Nullen dazu, ein Petabyte entspricht als 1.000.000.000.000.000 Bytes. Zum Vergleich: Für das Speichern von Rohtext auf einer durchschnittlichen DIN-A4-Seite benötigen Sie rund 1,8 Kilobyte an Speicherplatz.

WebText2 ist eine Erweiterung des ursprünglichen **OpenWebText-Korpus** mit allen Reddit-Posts zwischen 2005 und 2020. Reddit ist eine Social-News-Website, die in den USA weit verbreitet ist und auf der registrierte Benutzer Inhalte einstellen und anbieten können. Der

OpenWebText-Korpus gilt als Versuch, einen Datenbestand für das GPT-2-Modell von OpenAI zu schaffen, ohne urheberrechtlich geschützte Inhalte zu verwenden. Diese Einträge machen ein knappes Viertel der GPT-3-Trainingsdaten aus. Inhalte von OpenWebText2 umfassen Videos und Bilder, Links sowie Umfragen und Textbeiträge.

An dieser Stelle ist es wichtig zu betonen, dass Common Crawl und OpenWebText2 zusammen ganze 82 % der Trainingsdaten ausmachen, ohne dass diese Daten auf die Qualität der Inhalte geprüft worden wären. **Books1** und **Books2** sowie Wikipedia sind mutmaßlich bessere, weil abgesicherte Informationen, die vor der Veröffentlichung (oder danach) durch mehrere anspruchsvolle Qualitätsfilter laufen. Leeres Geschwurbel, die Fantasien von Verschwörungstheoretikern oder Behauptungen ohne fachliche Grundlage oder gesicherte Kompetenz findet man in diesen letzten 18 % wohl nicht. Die übrigen 82 % hingegen sind nur gesammelt und wahrscheinlich in qualitativer Hinsicht stark durchwachsen.

Wikipedia ist ohne Frage der hochwertigste Teil der Daten, macht trotz seiner gewaltigen Größe aber nur 3 % des Trainingsdatensatzes von GPT-3 aus! Die weltweit größte Enzyklopädie enthält mehr als 50 Millionen Artikel in vielen Sprachen, etwa sechs Millionen Artikel sind auf Englisch verfasst.



82 % der Trainingsdaten stammen aus nicht geprüften Quellen[1]

Weitere Datensätze, die in Sprachmodellen zum Einsatz kommen, sind z. B. **RealNews**, ein großer Korpus von Nachrichtenartikeln aus dem oben genannten Common-Crawl-Datensatz, aber begrenzt auf die

Nachrichtendomains, die von Google News indexiert werden. **Grover** ist aus RealNews entstanden und generiert Fake News zu Forschungszwecken. Ziel dabei ist natürlich, die Generierung neuronaler Desinformation besser erkennen und deren Verbreitung verhindern zu können. **CC-Stories** ist ein Datensatz für Common Sense Reasoning, der aus Aggregation von Dokumenten aus dem Common-Crawl-Datensatz erstellt wird.

Weitere Daten scheinen aktuell nicht verfügbar zu sein, außer vielleicht im Google Index, der aber aus verschiedenen Gründen wohl nicht zum Training der Sprachmodelle von OpenAI verwendet werden kann.

2.7 Ein Blick unter die Motorhaube: Wie funktioniert ChatGPT?

Seit den ersten Gehversuchen mit ChatGPT haben mich die folgenden Fragen zur Funktionsweise des Transformers beschäftigt und nicht mehr losgelassen:

- Wie kann die KI *lesen*, also die Bedeutung hinter Wörtern und Sätzen verstehen und Sprachintentionen daraus erkennen?
- Wie kann die KI *schreiben*, also korrekte und bedeutungsvolle Sätze aus Wörtern erzeugen, die auf meine Fragen eingehen?
- Wie kann die KI *Fakten* und *Zusammenhänge* aus Texten *extrahieren* und auf Anfrage neu zusammenstellen?
- Wie kann die KI *kreativ* und *problemorientiert kohärente Texte* erstellen?

Anfang der 90er-Jahre haben wir uns an der Uni leidenschaftlich mit Computerlinguistik beschäftigt, mit großer Hingabe Mustersätze in Baumstrukturen seziert, Lautketten in Phoneme und Morpheme zerlegt und das Für und Wider unterschiedlicher semantischer Modelle zur Beschreibung von Bedeutungen und deren Beziehungen diskutiert. Die *sprecherunabhängige Spracherkennung* war der Heilige Gral, schien unerreichbar für unsere Generation und ist heute eine Selbstverständlichkeit. Syntagmatische Verbindungen, also typische Abfolgen in Satzgefügen, waren in der Linguistik noch wenig erforscht, man kannte nur feste Wortverbindungen, Phraseologismen, Idiome oder Kollokationen. Die damalige Linguistik hatte mit Statistik noch nicht viel am Hut, Kollokationen waren eine der wenigen Ausnahmen, man definierte sie als eine Reihung von Wörtern, die häufiger gleichzeitig vorkommen als zufällig. Das Modell diente der Erklärung eines

sprachlichen Phänomens, kam aber über die Beispiele *eingefleischter Junggeselle* und *schütteres Haar* nicht wesentlich hinaus. Was war passiert und wie war ein solcher Quantensprung in der Verarbeitung natürlicher Sprache auf dem Computer in so kurzer Zeit möglich?

Man hatte ganz einfach die Methode geändert und das Personal ausgetauscht: Statt Linguisten immer neue Modelle bauen zu lassen und daraus Regeln für den Computer abzuleiten, hatten kluge Informatiker die Idee, der Maschine Texte zu geben und sie die Regeln einfach selbst finden zu lassen. Und anstatt diese Regeln dann in klassischen Wenn-dann-Strukturen zu speichern, hat man sie die Maschine selbst in tiefen neuronalen Netzen verankern und gewichten lassen. Anders ausgedrückt: Der Computer soll einfach so lernen, wie wir selbst die Sprache von unseren Eltern gelernt haben: mit viel Text, ohne explizite Regeln, dafür mit viel Lob und noch mehr Korrektur. Die Prinzipien des *Machine Learnings* führten offenbar sehr schnell zu vorzeigbaren Ergebnissen, hinzu kamen technische Fortschritte in der KI, eine massiv erweiterte Rechenleistung und die Verfügbarkeit riesiger Textmengen über das Internet. Die neue Herangehensweise und genannten Katalysatoren haben heutige Sprachmodelle à la GPT in einer damals unvorstellbar kurzen Zeit hervorgebracht.

Das grundlegende Arbeitsprinzip von GPT und anderen generativen KIs ist hinreichend bekannt: Der Transformer muss Texte auf sinnvolle Weise *Wort für Wort* ergänzen. Er sucht auf Basis aller aktuell vorliegenden Wörter (eigentlich Tokens) immer nach dem nächsten Wort und wählt dieses auf Basis von Wahrscheinlichkeiten aus einer langen Liste von Möglichkeiten. Ist das nächste Wort aus der Kandidatenliste ausgewählt, geht der gesamte Prozess wieder von vorn los, und so setzt der Transformer ganze Texte Wort für Wort auf Basis von Wortlisten und Wahrscheinlichkeiten zusammen. Schreiben ist für die KI also ein nicht endender Lückentext mit der Lücke stets am Ende!

Wenn ausschließlich Wahrscheinlichkeiten über das nächste Wort entscheiden, würde man einförmige Sätze mit geringer Varianz bekommen, und es würde niemals die Illusion natürlicher Sprache entstehen. Deshalb wählt die KI aus der Liste der möglichen Wörter eines mit niedrigerem Wahrscheinlichkeitsrang, damit das Ergebnis unserer Sprache ähnlicher wird. Welcher Rang ausgewählt wird, hat mit einem *Temperatur-Parameter* zu tun, der die Zufälligkeit der Reaktionen der KI bestimmt. Eine niedrige Temperatur erzeugt sehr zielgerichtete und vorhersehbare Ausgaben des Chatbots, je höher der Temperaturwert, umso vielfältiger und zufälliger werden die Texte.

Die Temperatur ist entscheidend, um Textsorte bzw. den Textstil anzupassen: Niedrige Temperaturen von 0,2 bis 0,5 erzeugen konservative und formelle Texte in der Art von Nachrichtenartikeln oder

wissenschaftlichen Abhandlungen. Das Modell wird tendenziell wahrscheinlichere und vorhersehbarere Wörter und Phrasen auswählen. Eine mittlere Temperatur bis etwa 1,0 erzeugt neutralen und allgemeinen Text. Informative Blogbeiträge und Produktbewertungen gelingen damit am besten, die Texte weisen eine ausgewogene Mischung aus Kreativität und Vorhersagbarkeit auf. Kreative Kurzgeschichten, Romane und Poesie, Werbung und Marketing sowie Brainstorming und Ideenfindung benötigen eine hohe Temperatur von etwa 1,2 bis 1,5 und führen zu kreativen und oftmals unerwarteten Texten. Mit einer Temperatur von 2,0 würden die generierten Texte wahrscheinlich weniger kohärent und weniger vorhersehbar sein. Sie könnten sich wiederholen, vom Thema abweichen oder ungewöhnliche Wortwahl und Satzstrukturen aufweisen. Es könnte auch zu mehr Unsinn oder weniger sinnvollen Inhalten führen.

Wie kommt ChatGPT eigentlich zu diesen Listen mit den Wahrscheinlichkeitswerten für das nächste Wort im Text? Dass es diese im Training aus Texten lernt, haben wir bereits gehört. Musste man diese Texte als Trainingsdaten vorher labeln oder taggen, wie wir im Beispiel mit den Bildern von Stühlen und Sesseln gelernt haben? Das wäre ein gigantischer Aufwand und eine Aufgabe für Generationen gewesen! Eine simple Idee spart diese Mühe und befähigt den Transformer zum unüberwachten Lernen: Im Text wird immer das nächste Wort abgedeckt (maskiert), der Computer muss es erraten. Dann wird geprüft, wie weit der Computer von der richtigen Lösung entfernt war und eine *Verlustfunktion* erstellt. Ziel des Lernprozesses ist es nun, das Ergebnis dieser Verlustfunktion immer weiter zu verkleinern und damit die Vorhersagen immer präziser werden zu lassen. Dazu werden im Wesentlichen die Gewichte im neuronalen Netz angepasst, um den mathematischen Verlust zu minimieren. Auch hier haben wieder Informatiker und nicht die Linguisten den Schlüssel zum Erfolg geliefert.

Wer noch tiefer in das Innenleben des Chatbots eintauchen möchte, dem sei Stephen Wolfram und »Das Geheimnis hinter ChatGPT« empfohlen, von dem auch die folgenden kenntnisreichen Zeilen stammen:

Bemerkenswert bei alldem ist am Ende, dass diese ganzen Operationen – so simpel sie für sich genommen sein mögen – es irgendwie gemeinsam schaffen, die »menschliche« Aufgabe des Generierens von Text so gut zu erledigen. Es muss noch einmal betont werden, dass es (zumindest soweit wir wissen) keinen »endgültigen theoretischen Grund« gibt, weshalb irgendetwas davon funktionieren sollte. Tatsächlich glaube ich, dass wir dies in unserer Diskussion als eine – möglicherweise überraschende – wissenschaftliche Entdeckung betrachten sollten: dass es in einem neuronalen Netz wie ChatGPT irgendwie möglich ist, die Essenz dessen zu erfassen, was

menschliche Gehirne beim Generieren von Sprache zustande bringen. (S. 87)

Ich glaube, der Erfolg von ChatGPT liefert uns den Beweis für ein grundlegendes und wichtiges Stück Wissenschaft: Er deutet an, dass es möglicherweise wichtige neue »Gesetze der Sprache« – und damit im Prinzip »Gesetze des Denkens« – zu entdecken gibt. In ChatGPT – das als neuronales Netz gebaut ist – drücken sich diese Gesetze bestenfalls implizit aus. Falls wir es schaffen könnten, diese Gesetze explizit zu machen, besteht das Potenzial, die Dinge, die ChatGPT macht, auf deutlich direktere, effizientere – und transparentere – Weise zu erledigen. (S. 98)

2.8Praktische Nutzung und KI-Prompting

In diesem Abschnitt soll es nun um die praktische Nutzung der KI gehen und darum, wie ChatGPT in verschiedenen Kontexten erfolgreich eingesetzt werden kann. Dazu möchte ich kurz auf die wirklich beachtlichen sprachlichen Fähigkeiten des Systems eingehen, bevor im folgenden Kapitel zahlreiche konkrete Beispiele für Prompts gegeben werden.

Der Hauptzweck von GPT liegt zwar in der Beantwortung von Fragen und der Bereitstellung von Informationen, das Modell kann aber deutlich mehr. Es kann eine Vielzahl von Textsorten in verschiedenen Stilen und für unterschiedliche Zwecke generieren und bearbeiten. Die folgenden Beispiele sind nur ein Auszug seiner Fähigkeiten: Artikel und Essays zu vielen wissenschaftlichen, technischen, kulturellen und anderen Themen. Kurzgeschichten und Gedichte, Skripte und Drehbücher für Theaterstücke, Filme oder Fernsehshows. Natürlich auch Dialoge zwischen zwei oder mehr Personen, Briefe und E-Mails von formell bis lässig. Rezensionen aller Art, ebenso Anleitungen und Tutorials zu einer Vielzahl von Problemfeldern. Sehr stark und erstaunlich einfallsreich ist das System auch im Bereich Werbung und Marketing, von Werbetexten bis hin zu Produktbeschreibungen und eingängigen Slogans textet das System so manche Agentur an die Wand. Und das zu einem Bruchteil der Kosten, fast ohne Wartezeiten oder Kapazitätsgrenzen! Ob man Programmcode als Untergruppe von Texten zählt oder nicht, der Bot beherrscht so viele Programmiersprachen, dass wohl kein menschlicher Entwickler mithalten kann: Python, Java und JavaScript, C, C+ + und C#, Ruby und natürlich PHP nebst HTML und CSS. Es erstellt Code in diesen Sprachen oder hilft bei der vertrackten Fehlersuche.

War je ein Schreiber so flexibel wie diese KI? ChatGPT kann für viele Zielgruppen schreiben, akademisch oder in der Sprache von Ämtern und Behörden, eher salopp für Jugendliche oder bildlich einfach für Kinder,

wobei Sie sogar ein beliebiges Alter vorgeben können. Der intelligente Chatbot versteht sich bestens auf Tonalität und Sentiment, kann überzeugend, informativ, formell oder locker schreiben und in den Texten bestimmte Stimmungen wiedergeben, also positiv, negativ, vorsichtig, zuversichtlich oder deprimiert schreiben. Vokabular und Ausdrucksweise können den Jargon und den speziellen Sprachstil einer beliebigen Gruppe nachahmen, vom Rocker bis zum Pfarrer, vom Polizisten bis zum Oberstudiendirektor, vom Juristen bis zum Poeten.

Natürlich ist das Sprachmodell ein echter Meister in allen rein sprachbezogenen Aufgaben: Zusammenfassungen aller Art von Büchern, Artikeln, Vorträgen in beliebiger Länge (einfach die gewünschte Wortanzahl vorgeben). Umschreiben und Paraphrasieren, also einen beliebigen Inhalt mit anderen Worten nochmals wiedergeben. Korrektur und Lektorat verbessern Form und Stil beliebiger Texte und erklären auf Wunsch jeden einzelnen Fehler wie ein unermüdlicher Lehrer im Deutsch- bzw. Sprachunterricht. Als Lehrer für Fremdsprachen oder als Tutor im Sprachunterricht hält das System erstaunliche Fähigkeiten für Lehrende und Lernende bereit. Die bemerkenswerte Mehrsprachigkeit kann also nicht nur für maschinelle Übersetzungen, sondern auch zum Lernen vieler Sprachen genutzt werden! Wenn Sie die Qualität unterschiedlicher KIs zur maschinellen Übersetzung selbst testen und vergleichen wollen, nehmen Sie einen beliebigen Text, übersetzen Sie diesen von Sprache A nach B, dann von B nach C und schließlich von C wieder zurück in die Ausgangssprache.

Prompts sollten möglichst klar und präzise formuliert sein ohne unnötige Wörter oder leere Floskeln. Prompts in einer klar erkennbaren Sprache werden in derselben Sprache beantwortet, obwohl viele Beobachter vermuten, dass die KI ausschließlich auf Englisch denkt (also alle Eingaben ins Englische übersetzt, verarbeitet und das Ergebnis dann in die Zielsprache zurückübersetzt). Fragen Sie immer so genau wie möglich. Wenn Sie fünf Tipps von der KI haben wollen, dann nennen Sie diese Zahl auch im Prompt, statt nur nach *Tipps* zu fragen. Korrekte Grammatik und Orthografie sollten eine Selbstverständlichkeit sein, damit es nicht zu Missverständnissen zwischen der KI und dem Prompter kommt. Aktive Satzkonstruktionen sind meistens kürzer und prägnanter als die Verwendung des Passivs. Geben Sie den Kontext an, wenn er zum Lösen der Aufgabe notwendig ist. Golf kann man *spielen*, in einem Golf kann man *fahren* oder an einen Golf *reisen*. Je mehr Hintergrundinfos die KI hat, umso spezifischer kann der Output erfolgen. Vermeiden Sie Unverständliches wie Abkürzungen, Umgangssprache oder Slang bzw. doppeldeutige Wörter.

Und wie lang darf so ein Prompt sein? Feste Grenzen gibt es nicht, außer der des Kontextfensters. Was man dazu wissen sollte: In ChatGPT

wird die Tokengröße zwischen Prompt und Ausgabe geteilt, je länger also der Prompt, desto kürzer ist (theoretisch) die Ausgabe. Manche Prompts lassen sich logisch sehr gut teilen, indem Sie z. B. eine Zusammenfassung eines Textes im ersten Prompt anfordern und dann den Text in einem zweiten Prompt folgen lassen.

Prompt-Engineering? Schon mal gehört? Das ist die Kunst oder die Wissenschaft, eine KI mit den optimalen Anweisungen bzw. Inputs zu versorgen. Klingt simpel, ist es auch, aber die Besten in dieser Disziplin können sich aktuell den Arbeitsplatz aussuchen und bekommen Traumgehälter gezahlt. Wenn Ihr Job von der KI bedroht ist, Prompt-Entwickler sind gesucht.

2.9Plugins, Erweiterungen und KI-Schreibassistenten

Was ChatGPT auch in der aktuellen Version noch nicht kann, steuert vielleicht eine der vielen Systemerweiterungen bzw. eines der **Plugins** für den Browser bei. Wer die KI zu einem aktuellen Ereignis befragt, scheitert zunächst am Wissenstand des Bots, dessen Trainingsdaten nur bis zum September 2021 reichen. Befragt man die KI zum Krieg in der Ukraine, kann diese nur über einen bewaffneten Konflikt mit prorussischen Separatisten und russischen Streitkräften in den Regionen Donezk und Luhansk im Osten des Landes berichten. Installiert man nun aber eine der zahlreichen Erweiterungen, die auf aktuelle Suchergebnisse von Google & Co zugreifen und neue Fakten integrieren können, lautet die Antwort deutlich aktueller: *Der Krieg in der Ukraine ist ein ernstes und komplexes Thema, dessen Dynamik sich ständig verändert.* Die meisten Browser-Erweiterungen sind kostenlos für Chrome und Firefox sowie auf den Plattformen Windows, Apple und Linux verfügbar.

WebChatGPT erweitert alle ChatGPT-Antworten mit tagesaktuellen Web-Suchergebnissen. Dazu aktiviert man einfach den Abruf einer Suchmaschine mit dem Schalter *Web access* (Webzugriff). Stellt man dem Chatbot mit aktiviertem Webzugriff nun eine offensichtlich aktuelle Frage, sprudeln die Ergebnisse nur so aus der KI heraus: Am Ende erhält man eine topaktuelle, zusammenfassende Antwort inklusive Angabe der Websites, die dafür befragt und ausgewertet wurden.

Welche Suchmaschine verwendet wird, ist ebenso einstellbar wie die Suchtiefe, also ob ich eher schnell die ersten Absätze lesen oder sehr gründlich die gesamten Inhalte einer Quelle auswerten möchte. Wie viele Suchergebnisse werden berücksichtigt? Die Voreinstellung im Firefox ist mit *3 results* für die meisten Fragen bestimmt ausreichend, lässt sich aber schrittweise auf zehn Suchergebnisse erweitern oder mit

max results bis auf die Spitze treiben. Wer das Suchergebnis nicht dem Zufall und irgendeiner Website überlassen möchte, kann die Operatoren */site* oder */page* nutzen: Ersterer beschränkt die Suche auf eine Domain, z. B. */site:wikipedia.com*, letzterer gar auf eine einzelne Seite, z. B. */page:blog.google/intl/de-de/unternehmen/technologie/bard-deutschland-verfuegbarkeit-ki/*. Mit dieser Funktion wird Ihnen ChatGPT jeden beliebigen Blogartikel in gewünschter Länge zusammenfassen. Die Option *One-click ChatGPT-Prompts* (ChatGPT-Eingabeaufforderungsbibliothek) bietet eine kostenlose Sammlung mit Hunderten von Prompts, die bei der Lösung von alltäglichen Problemen im privaten Bereich und sich wiederholenden Aufgaben im Unternehmen helfen können. Auf Wunsch sammeln und verwalten Sie damit auch die eigenen KI-Prompts.

Das Thema Import und Export behandeln die folgenden beiden Plugins aus entgegengesetzten Richtungen: Ask Your PDF ist eine Erweiterung, die Informationen aus PDF-Dateien extrahiert und für ChatGPT bereitstellt bzw. dessen Inhalt in den Chatbot hochlädt. Der ChatGPT Prompt Genius macht genau das Gegenteil: Über den Menüpunkt *Share & Export* exportiert die Erweiterung ganze Chat-Verläufe als PDF oder Bilddatei. Beide Plugins sind keine bahnbrechenden Erfindungen, sondern nützliche kleine Helfer, die den Alltag erleichtern. Wer sich ausschließlich in der Google-Welt bewegt, den Chrome-Browser und Google Mail nutzt, dem kann der ChatGPT Writer automatische Antworten auf E-Mails generieren. Der Tweet Hunter und Copy.ai formulieren Beiträge, Blogposts, Newsletter und Anzeigentexte. Irgendwann in einer nicht allzu fernen Zukunft werden Maschinen dann mit Maschinen kommunizieren – bleibt nur zu hoffen, dass wir Menschen die gewonnene Zeit dann wenigstens sinnvoll nutzen!

Plugins ergänzen die Oberfläche von ChatGPT und erweitern seine Möglichkeiten. Einen ganz anderen Weg gehen Programme mit eigener Oberfläche, bei denen ein Sprachmodell im Hintergrund zuarbeitet. Nennen wir sie **KI-gesteuerte Schreibmaschinen**, die auf die Bedürfnisse und den Workflow von professionellen Schreibern ausgerichtet sind. Diese eigenständigen Software-Tools nutzen künstliche Intelligenz im Hintergrund, um Benutzern beim Schreiben zu helfen mit Unterstützungen von der einfachen Rechtschreib- und Grammatikprüfung bis hin zur Erstellung ganzer Textabschnitte oder sogar kompletter Artikel. Diese Tools unterscheiden sich in ihrer Spezialisierung und möglichen Anwendungsbereichen und haben unterschiedliche Stärken und Schwächen. Sie sind besonders beliebt bei Journalisten, Copywritern, Bloggern und anderen Content-Vermarktern. Selbst angesehene Schriftsteller und Autoren greifen mittlerweile mehr oder weniger offen zu diesen Hilfsmitteln. Viele zählen auch ChatGPT zu

den KI-gesteuerten Schreibmaschinen, wir wollen hier aber andere Tools und ihre unterschiedlichen Ansätze vorstellen.

DeepL Write ist ein KI-gesteuerter Schreibassistent, der *Bessere Texte im Handumdrehen* verspricht und tatsächlich die Qualität beliebiger Texte verbessern kann. Er soll klare, präzise und fehlerfreie Texte ermöglichen und die schriftliche Kommunikation auf ein neues Niveau heben. Das Tool verwendet ein eigenes Sprachmodell, ob DeepL Write auf dem gleichen KI-Modell wie der Übersetzungsdienst von *DeepL* basiert oder ob es ein speziell dafür entwickeltes Modell gibt, ist leider nicht klar dokumentiert. Die Oberfläche der Webanwendung *deepl.com/de/write* arbeitet mit zwei Fenstern, links muss der Anwender seinen Text eingeben, um im rechten Fenster Verbesserungsvorschläge zu erhalten. Wer ein Konto einrichtet und sogar den Pro-Account nutzt, bekommt mehrere Alternativen zu jeder Korrektur und kann diese im rechten Fenster mit der Maus übernehmen oder verwerfen. DeepL Write glänzt mit präzisen Korrekturen und sinnvollen Formulierungsvorschlägen.

LanguageTool stellt sich dem Benutzer als »Schreibassistent« vor, der Texte überprüfen und umformulieren kann. LanguageTool ist ein mehrsprachiges Online-Tool *languagetool.org/de*, das nach Erstellung eines kostenlosen Accounts ablenkungsfreies Schreiben, ein persönliches Wörterbuch und ein Punktesystem zur Bewertung eigener Texte anbietet. Wer in Sachen Stil, Wortwahl und Sprachqualität das Maximum herausholen möchte, wählt den »akribischen Modus« von LanguageTool, um noch mehr Stil- und Formulierungsvorschläge auch bei langen und komplexen Sätzen zu erhalten, Umgangssprache und Redundanzen erkennen zu lassen und Synonyme zur Vermeidung von Wiederholungen aufgelistet zu bekommen. Ein spontaner Test mit drei langen Sätzen ergab nur wenig brauchbare Vorschläge, dafür wurden alle Rechtschreibfehler präzise erkannt und sinnvolle Alternativen zur Korrektur präsentiert.

Grammarly ist ein weiteres beliebtes Tool, das ebenfalls auf Grammatikprüfung und Stilverbesserung fokussiert ist. Weitere Tools wie **Quillbot** sind auf die Paraphrasierung und Textüberarbeitung für die akademische Welt spezialisiert. **ShortlyAI** ist ein KI-gesteuertes Texterstellungstool, das im Gegensatz zu seinem eigenen Namen auf die Generierung von ausführlichen Blogposts und langen Artikeln spezialisiert ist. **Textio** und **Jasper** sind auf das geschäftliche Umfeld ausgerichtet und konzentrieren sich auf ganz bestimmte Spezialaufgaben: Textio will in Stellenanzeigen die besten Kandidaten ansprechen, Jasper ist ein Tool für Content-Marketing und SEO, mit dem Marketingverantwortliche Inhalte für verschiedene Plattformen generieren können. Wie bei allen Tools ist die Qualität der generierten Texte variabel und erfordert oft menschliche Nacharbeit.

2.10Maschinell erstellte Texte erkennen: AI-Detektoren

Im Rahmen eines Google Hangouts wurde von einem Anwender die Frage gestellt, wie die Google-Suche KI-generierte Texte einordnen würde. John Mueller, Senior Search Analyst bei Google und erste Instanz für alle SEOs und Webmaster, beantwortete diese Frage damit, dass alle computergenerierten Texte laut den Webmaster-Guidelines grundsätzlich als *Spam* anzusehen sind. Dabei sei es unerheblich, ob eine KI zum Einsatz gekommen ist oder nicht. »Für uns würden solche Texte im Wesentlichen immer noch in die Kategorie der automatisch generierten Inhalte fallen, die wir fast von Anfang an in den Richtlinien für Webmaster hatten«, sagte Mueller den Teilnehmern. Webmaster fürchten den Ausdruck *Spam*, denn wenn ihn Google auf der eigenen Website mehrfach entdeckt, kann das eine empfindliche Google-Strafe nach sich ziehen. Google-Strafen bedeuten immer drastische Rankingverluste und im schlimmsten Falle sogar das komplette Verschwinden einer ganzen Domain aus dem Google-Index.

Ob Google KI-generierte Texte denn überhaupt automatisch erkennen könne, war die naheliegende nächste Frage. Mueller wollte dies nicht ausdrücklich bejahen, stellte aber mögliche Konsequenzen dar: »Das kann ich nicht behaupten. Aber wenn wir sehen, dass etwas automatisch generiert wird, kann das Webspam-Team definitiv etwas dagegen unternehmen.«

Ähnliches droht jetzt auch Studenten beim Anfertigen schriftlicher Hausarbeiten, die Verwendung einer KI ist strengstens verboten. Werden Studenten dabei erwischt, droht auch hier mit der Exmatrikulation die Höchststrafe. Die Verwendung von ChatGPT gilt als Schummelei und wird mit maximaler Härte geahndet. Doch wie erkennt Google oder der kritische Professor eigentlich, ob ein vorliegender Text mit KI geschrieben oder von Menschenhand erstellt wurde? Das ist nicht ganz einfach und eigentlich fast unmöglich, dennoch gibt es Wege, die zumindest einen Verdacht erhärten können.

Ein experimentierfreudiger Professor soll mit einer Gruppe von Studenten im Rahmen einer Challenge versucht haben, aus einer Reihe von manuell gefertigten Hausarbeiten diejenige zu identifizieren, die mit ChatGPT erstellt wurde. Und angeblich fand er sie auch tatsächlich, weil nur eine einzige Arbeit keine orthografischen Fehler enthielt. In der Vorbereitung des vorliegenden Buches sind die Autoren auf das Buch »Künstliche Intelligenz ChatGPT: Die Sprachrevolution« von Adam Jones gestoßen, das auf eine merkwürdige Weise inkonsistent war. Inhaltsleere Passagen voller formaler Fehler wechselten sich ab mit solchen, die fehlerfreie und sinnvolle, kohärente Inhalte enthielten. Nach wenigen

Kapiteln war ziemlich klar, dass der Autor offenbar lange Passagen aus ChatGPT notdürftig mit selbst geschriebenen Einleitungen und Überleitungen verknüpft hatte. Wer so nachlässig mit dem Bot umgeht, ist leicht zu entlarven. Umgekehrt gilt aber auch: Wer geschickt genug mit dem Tool arbeitet und gut tarnt, dem kann die Verwendung der KI-Chatbots nicht sicher nachgewiesen werden. Der Wunsch, KI-generierte Texte als solche zu erkennen, hat längst eine ganze Reihe von Tools hervorgebracht: die sogenannten *KI-Detektoren*.

Plagiatsprüfer wie *duplichecker.com*, *scribbr.de* und *copyleaks.com* sind weit verbreitet und leisten gute Arbeit. Eine halbe Doktorarbeit einfach nur abzuschreiben, wagt heute niemand mehr. Wer die KI geschickt nutzt, um Texte mehrfach umformulieren zu lassen, bekommt bei jeder automatisierten Plagiatsprüfung schnell grünes Licht. Karl-Theodor zu Guttenberg, dem 2011 wegen nachgewiesener »Urheberrechtsverletzungen« der Doktorgrad aberkannt wurde, hätte heute leichtes Spiel, Plagiate wirkungsvoll zu tarnen. Auch die Hoffnung, ChatGPT selbst könnte sich an Texte »erinnern«, die es selbst erstellt hat, erweist sich als Irrglauben. Der Bot selbst sagt dazu:

```
Ich kann keine spezifischen Texte erkennen oder speichern, die in
früheren Sitzungen von mir oder einer anderen Instanz von ChatGPT
produziert wurden. Jede Anfrage, die an mich gestellt wird, wird
als isoliert betrachtet, und ich behalte keine Erinnerung an
frühere Anfragen oder Antworten. Dies ist wichtig, um die
Privatsphäre der Benutzer zu schützen.
```

Einen der ersten Detektoren veröffentliche der Hersteller sogar selbst, den *OpenAI AI Text Classifier*. Andere Tools sind GPTZero, Orginality.ai und DetectGPT nutzen teilweise dieselbe KI-Technologie.

Wenn man ein neuronales Netz zur Klassifizierung in einem Detektor verwendet, trainiert man es mit KI-generierten Texten in der Hoffnung, Deep Learning würde relevante Muster in KI-Texten erkennen und damit die Spreu sicher vom Weizen handgeschriebener Texte trennen. In der Praxis erweisen sich aber alle Detektoren als mehr oder weniger zuverlässig und bezichtigen teilweise sogar von Hand geschriebene Texte, sie seien von einer KI produziert.

Die Methoden der einzelnen Tools sind unterschiedlich: **GPTZero** berechnet aus einer beliebigen Textprobe die sogenannte *Perplexity*, also die statistische Vorhersagbarkeit des nächsten Wortes. Dieses Konzept wird auch verwendet, um die Güte eines Sprachmodells zu testen – je weniger vorhersagbar (also überraschender) das nächste Wort ist, umso besser ist das Modell. Die Perplexity eines von Hand geschriebenen Textes ist (zumindest in der Theorie) höher als die eines Sprachmodells.

Ein anderes Prinzip der Detektoren lautet »An ihren Fehlern werdet ihr sie erkennen«. ChatGPT produziert nämlich systematisch kleine Fehler, die mittlerweile sogar in speziellen Listen bekannt gemacht werden und anhand derer man dann den Nachweis führen kann. Andere Tools brauchen zur Entdeckung des Urhebers den Zugriff auf das bei der Erstellung des Textes verwendete Tool inklusive einer Versionsangabe. Aber wenn jemand diese Information freiwillig unter seinen Text schreibt, wozu braucht es dann noch einen Detektor?

Für alle Tools gilt: Ein negatives Prüfergebnis besagt nur, dass das Tool keine Hinweise auf seinen künstlichen Ursprung gefunden hat. Zur händischen Verschleierung des Urhebers eines computergenerierten Textes gibt es schon jetzt viele Tipps, die fast alle sehr gut funktionieren, schon weil mit jeder manuellen Änderung die »Handschrift« eines Autors einfließt und das Risiko der Enttarnung verkleinert wird.

Ein Gedanke erscheint mir abschließend noch wichtig und nimmt eigentlich allen Detektoren noch im Hafen schon den größten Wind aus den Segeln: Wenn ein Algorithmus jemals in der Lage sein sollte, zuverlässig zu erkennen, ob ein Text von Hand oder von einem Bot erstellt wurde, dann gibt es kurze Zeit später ein neues Programm, das die Texte wieder erfolgreich gegen den Detektor tarnt. Hier schlägt ein Algorithmus den anderen! Der Wettlauf zwischen automatischer Tarnung und Entdeckung erinnert ein wenig an den Wettbewerb von Ärzten, die auf der einen Seite beim Doping immer neue Methoden der Tarnung entwickeln, und ihre Kollegen auf der anderen Seite, die die Tests laufend verfeinern. Anders als beim Doping sind kluge KI-Sünder aber praktisch nicht zu entdecken. Positive Dopingtests sind gerichtsfeste Beweise gegen die Sünder, Befunde von KI-Detektoren nur vage Indizien.

2.11 Bard: Das KI-Experiment von Google

»Google-Aktie verliert 100 Milliarden Dollar«, »Googles Chatbot Bard patzt in der ersten Präsentation« und »Googles Chatbot debütiert mit einer Falschinformation« titelten Anfang des Jahres die Fachmagazine. Was war passiert? Der Grund war eine fehlerhafte Antwort des Chatbots von Google bei der ersten Präsentation. Auf die Frage, von welchen Entdeckungen des James-Webb-Weltraumteleskops man einem 9-jährigen Kind erzählen könnte, gab Bard eine Antwort mit einem groben Schnitzer: Das Teleskop hätte das allererste Foto eines Planeten außerhalb unseres Sonnensystems gemacht. Was schlichtweg falsch ist! Die erste Aufnahme eines Exoplaneten wurde nämlich schon 2004 vom europäischen Very Large Telescope (VLT) gemacht. Autsch!

Zum Hintergrund: Nach der Veröffentlichung von ChatGPT Ende 2022 erkannte Google, der Chatbot könnte der Suchmaschine erhebliche Marktanteile streitig machen. Auch weil der Erzrivale im Suchmaschinengeschäft *Bing* heißt und von Microsoft stammt, das wiederum an OpenAI beteiligt ist. Google hatte intern auch schon den sogenannten *Code Red* ausgegeben, was nichts weniger bedeutet, als dass man das eigene Geschäftsmodell ernsthaft bedroht sieht oder gar die komplette Existenz von Google auf dem Spiel steht. CEO Sundar Pichai beschloss, mit einer Gegenattacke zu antworten und den hauseigenen Chatbot Bard vorzustellen, der nach seiner Fertigstellung die eigene Websuche erweitern sollte. Das war nicht nur überraschend, sondern auch erkennbar überhastet und noch dazu völlig verpfuscht.

Die Veröffentlichung von ChatGPT hatte nach der ersten Begeisterung die Erkenntnis ausgelöst, dass der Bot zwar fast jede Frage beantwortet, bei den Fakten aber manchmal meilenweit daneben liegt. »Gut, dass es Google gibt«, sagten oder dachten viele, »die können das mit den Fakten!« Und dann liefert das Google-Tool schon bei der ersten Frage einen so peinlichen Fakten-Fehler! Das Tragischste von allem aber war, dass die Frage an Bard nicht etwa zufällig auf einer Pressekonferenz gestellt wurde, Google hat das Thema Weltraumteleskop wohl in einem Blogpost selbst als Beispiel zur Demonstration für Bard ausgewählt und den Fehler in der Antwort schlichtweg übersehen. Oh nein! Die Google-Aktie sank am Tag nach der Vorstellung auf Twitter um 7,8 Prozentpunkte. Seitdem ist es still um Bard geworden, aber testen Sie den Bot doch einfach mal selbst unter bard.google.com. Fragt man Bard, was es denn besser könne als ChatGPT, bekommt man folgende Antwort:

```
Beantwortung von Fragen auf informative Weise: Bard ist besser
darin, Fragen auf informative Weise zu beantworten, auch wenn sie
offen, herausfordernd oder seltsam sind. Bard kann beispielsweise
Fragen zu aktuellen Ereignissen, wissenschaftlichen Themen oder
literarischen Werken beantworten.
Ich bin in Bezug auf die Korrektheit von Fakten besser als
ChatGPT. Ich bin auf einem riesigen Datensatz aus Text und Code
trainiert, der aus einer Vielzahl von Quellen stammt, darunter
Bücher, Artikel, Websites und andere öffentliche Dokumente.
Dadurch habe ich einen besseren Überblick über die Welt und kann
Fakten genauer darstellen.
In einer Studie von Stanford University wurde festgestellt, dass
Bard bei der Beantwortung von Fragen zu aktuellen Ereignissen,
wissenschaftlichen Themen und literarischen Werken genauer ist
als ChatGPT. Bard machte nur 2 % Fehler, während ChatGPT 10 %
Fehler machte.
Ich bin jedoch noch in der Entwicklung und lerne ständig dazu. Es
```

```
ist möglich, dass ChatGPT in Zukunft in Bezug auf die Korrektheit
von Fakten besser wird.
```

Weitere Mitbewerber um den Transformer-Thron: *Jasper Chat* gehört auf jeden Fall zu den guten Textgeneratoren auf dem Markt und wird zu den besten Alternativen zu ChatGPT gezählt. *Neuroflash* kann neben Text auch Bilder erzeugen, die magische Feder – so nennt sich der Texteditor – hat weltweit bereits über 500.000 registrierte Nutzer. Die KI-Modelle im Hintergrund der genannten Tools heißen allerdings GPT, nur Bard baut auf ein eigenes Modell mit eigenen Trainingsdaten.

3Prompts für Schule, Beruf und Freizeit

Hier nun also der praktische Teil des Buches: Eine Sammlung von Prompts aus zwölf ganz unterschiedlichen Bereichen für Schule, Beruf und Freizeit. Wir haben uns für diese Form entschieden, weil Ihnen die folgenden Beispiele am ehesten einen Eindruck geben, was das KI-Modell alles für Sie tun kann. Probieren Sie es doch selbst aus!

3.1Hilfe in Schulfächern, Hausaufgaben, Prüfungsvorbereitung

> >Erkläre mir ein Thema aus der Algebra: Wie löst man quadratische Gleichungen mittels der quadratischen Ergänzung und gib ein praktisches Beispiel.

> >Erstelle eine Zeitleiste für das Fach Geschichte mit den wichtigsten Ereignissen des Zweiten Weltkriegs zwischen 1939 und 1941.

> >Hilf mir im Fach Biologie: Beschreibe den Prozess der Fotosynthese in Pflanzenzellen. Welche Rolle spielen Chloroplasten dabei?

> >Analysiere das Thema ’Einsamkeit’ im Roman ’Von Mäusen und Menschen’ von John Steinbeck für das Fach Englisch.

> >Erkläre das newtonsche Gesetz der Gravitation im Fach Physik und wie es die Bewegung der Planeten um die Sonne beeinflusst.

> >Beschreibe die Eigenschaften von Alkalimetallen für das Fach Chemie und warum sie in der ersten Gruppe des Periodensystems stehen.

> >Ich brauche Hilfe im Fach Geografie: Skizziere die verschiedenen Klimazonen der Erde und erkläre die Hauptmerkmale jeder Zone.

> >Beschreibe den Übergang von der Renaissance zum Barock in Europa für das Fach Kunst bzw. Kunstgeschichte. Welche Künstler und Werke sind bezeichnend für diese Zeiträume?

> >Ich möchte mich auf eine Prüfung vorbereiten. Gib mir fünf effektive Techniken zur Vorbereitung auf eine schriftliche Prüfung und erkläre mir, wie man den Lernstoff effizient strukturiert.

> >Mein Kind hat Probleme mit Sachaufgaben. Bitte entwirf zehn Sachaufgaben für eine 12-jährige Gymnasiastin, deren Anspruch sich

kontinuierlich steigert. Lösungen brauche ich auch.

> >Erkläre den Unterschied zwischen Genotyp und Phänotyp im Fach Biologie. Wie beeinflusst die Mendelsche Vererbung die Ausprägung von Merkmalen in nachfolgenden Generationen?

> >Analysiere für das Fach Deutsch die Rolle des Individuums in der Gesellschaft in Franz Kafkas 'Die Verwandlung'. Welche literarischen Mittel verwendet Kafka, um Isolation und Entfremdung darzustellen?

> >Beschreibe im Fach Physik, wie ein Transformator funktioniert. Welche physikalischen Prinzipien liegen seiner Funktion zugrunde?

> >Erkläre die Struktur und Eigenschaften von Alkanen für das Fach Chemie. Wie unterscheiden sie sich von Alkenen und Alkinen?

> >Beschreibe die Unterschiede zwischen den musikalischen Epochen Barock und Klassik in der Musik. Welche Komponisten und Werke sind für diese Zeiträume prägend?

3.2Fremdsprachen lernen für Urlaub, Schule und Beruf

> >Stelle eine Liste mit 20 nützlichen Redewendungen für eine Reise nach Frankreich zusammen und gib jeweils einen Beispielkontext an.

> >Erstelle einen Leitfaden mit den 50 wichtigsten Sätzen, die man braucht, um sich in Portugal zurechtzufinden, von der Begrüßung bis hin zu Notfallsituationen.

> >Hilf mir bei der sprachlichen Etikette im geschäftlichen Umfeld: Beschreibe die kulturellen und sprachlichen Besonderheiten bei Geschäftstreffen in Japan. Welche Höflichkeitsformen und speziellen Ausdrücke sollte man kennen?

> >Erkläre den Unterschied zwischen den spanischen Verben 'ser' und 'estar'. Wann verwendet man welches Verb, und welche Beispielsätze gibt es?

> >Beschreibe, wie Humor in der britischen Kultur funktioniert und welche sprachlichen Besonderheiten es gibt, die Nicht-Muttersprachlern oft entgehen.

> >Bewerte die drei besten Lernprogramme bzw. Apps zum Erlernen der türkischen Sprache in Bezug auf Benutzerfreundlichkeit, Inhalt und Effektivität.

> >Erstelle einen 7-tägigen Plan für jemanden, der in der Vorbereitung auf einen Urlaub in Portugal täglich 20 Minuten Portugiesisch lernen möchte. Welche Themen und Vokabeln sind besonders relevant?

> >Ich möchte mich auf eine mündliche Prüfung im Fach Spanisch vorbereiten: Simuliere ein 5-minütiges Gespräch zwischen einem Interviewer und einem Kandidaten in einer mündlichen Spanischprüfung. Welche Fragen könnten gestellt werden und wie könnte ich darauf antworten?

> >Hier sind fünf Sätze auf Französisch mit grammatikalischen Fehlern. Kannst du die Fehler identifizieren und korrigieren?

> >Wie verfasst man eine formelle E-Mail auf Chinesisch an einen Geschäftspartner? Welche Redewendungen und Höflichkeitsformeln sollten berücksichtigt werden?

> >Welche Techniken können helfen, das Hörverständnis im Italienischen für Fortgeschrittene zu verbessern, insbesondere, wenn es um regionale Dialekte geht?

> >Entwickle ein einfaches Brettspiel für Englischlerner der 10. Klasse Gymnasium, das sowohl den Wortschatz der Schüler erweitert als auch Spaß macht.

3.3Bewerbung, Lebenslauf und Fragen im Vorstellungsgespräch

> >Entwirf ein Anschreiben für einen Softwareentwickler mit fünf Jahren Erfahrung, der sich für eine Stelle in einem innovativen Start-up im Bereich erneuerbare Energien bewirbt.

> >Wie erstelle ich ein modernes Design für einen Lebenslauf, das sowohl professionell als auch kreativ ist. Worauf muss ich besonders achten und was darf ich nicht tun?

> >Formuliere drei Karriereziele für einen Absolventen im Bereich Umweltwissenschaften, der in einer NGO arbeiten möchte, die sich mit dem Klimawandel befasst.

> >Welche Fragen könnte ein Interviewer im Vorstellungsgespräch stellen, um herauszufinden, wie ein Bewerber in stressigen Situationen oder unter Druck reagiert?

> >Wie könnte eine Person mit mehrjähriger Erfahrung in

Projektmanagement eine Frage nach seinem größten beruflichen Erfolg beantworten?

> >Nenne fünf Soft Skills für einen Lebenslauf, die für einen Kundendienstmitarbeiter besonders wichtig wären, und beschreibe kurz, warum sie relevant sind.

> >Nenne häufige Fehler, die Menschen in ihren Lebensläufen machen, und gib Tipps, wie sie vermieden werden können.

> >Gib Ratschläge und Vorschläge, wie jemand, der zehn Jahre in der Finanzbranche gearbeitet hat, seinen Lebenslauf neu gestalten kann, um in den Non-Profit-Sektor zu wechseln.

> >Schlage einige Fragen vor, die ein Bewerber stellen könnte, um mehr über die Unternehmenskultur und Teamdynamik des potenziellen neuen Arbeitgebers zu erfahren.

> >Verfasse ein Anschreiben für einen Lehrer, der nach 20 Jahren im Bildungswesen eine Stelle im Personalwesen eines großen Technologieunternehmens übernehmen möchte.

> >Zeige, wie jemand seine Führungsqualitäten in einem Lebenslauf hervorheben kann, nachdem er erfolgreich ein Team von 20 Personen geleitet hat.

> >Welche Strategie könnte ein Bewerber anwenden, um während eines Vorstellungsgesprächs ein höheres Gehalt und zusätzliche Vergünstigungen zu verhandeln?

> >Wie könnte ich auf die Frage 'Warum haben Sie in den letzten sechs Monaten keine Anstellung gefunden?' im Vorstellungsgespräch höflich und positiv antworten?

> >Ich habe eine berufliche Lücke im Lebenslauf. Wie kann ich eine Auszeit von einem Jahr im Lebenslauf positiv darstellen, die für eine Weltreise genutzt wurde.

3.4Planung und Durchführung von Veranstaltungen und Events

> >Beschreibe eine detaillierte Idee für eine 'Zurück in die 80er'-Themenparty, inklusive Musikvorschlägen, Dekorationen und passenden Snacks.

> >Erstelle Struktur und Zeitplan für eine 2-tägige digitale

Marketingkonferenz und berücksichtige Keynote-Redner, Workshops und Networking-Sessions.

> >Gib eine Liste von kreativen Ideen für Outdoor-Hochzeiten im Herbst, von der Dekoration bis zu wetterbedingten Notfallplänen.

> >Skizziere ein Konzept für ein lokales Kulturfestival, das traditionelle Musik, Tänze und Speisen einer bestimmten Kultur oder Region präsentiert.

> >Entwirf einen Plan für die Einführung eines neuen, umweltfreundlichen Produkts. Berücksichtige dabei sowohl physische Präsentationen als auch digitale Werbemaßnahmen.

> >Liste fünf kreative Teambuilding-Aktivitäten auf, die ein Team von Steuerfachangestellten während eines Firmen-Retreats durchführen kann, um die Zusammenarbeit zu fördern.

> >Stelle eine Schritt-für-Schritt-Anleitung für die Planung und Durchführung einer Kunstausstellung in einer lokalen Galerie zusammen.

> >Skizziere einen Plan für ein 3-tägiges Rockmusik-Festival, inklusive Bühnenaufbau, Künstlerlogistik und Sicherheitsprotokollen.

> >Beschreibe ein Konzept für ein Charity-Dinner, das Spendengelder für einen wohltätigen Zweck sammeln soll, inklusive Menüauswahl, Unterhaltung und Spendenstrategien.

> >Entwirf eine Idee für ein Pop-up-Restaurant in einem städtischen Park, das regionale und saisonale Speisen in einer einzigartigen Umgebung präsentiert.

> >Stelle ein detailliertes Konzept für einen interaktiven Kochworkshop zusammen. Berücksichtige dabei die notwendigen Zutaten, Küchenutensilien und die Schritte zur Zubereitung eines dreigängigen Menüs.

> >Skizziere eine Agenda für das 25-jährige Firmenjubiläum eines Automobilzulieferers, das in einem eleganten Ballsaal stattfindet. Plane dabei eine Begrüßungsrede, eine Multimedia-Präsentation zur Unternehmensgeschichte und ein festliches Menü in mindestens fünf Gängen mit Optionen für Veganer.

> >Entwirf eine Piraten-Themenparty für einen Kindergeburtstag mit Kindern im Alter von sechs bis acht Jahren. Überlege dir Spiele, Dekorationen und Snacks, die ins Thema passen.

> >Skizziere die Planung von einem ’Tag der sauberen Stadt’ mit

Müllsammelaktionen, Workshops zum Thema Recycling und Infoständen zu umweltfreundlichem Verhalten. Gib mir fünf Vorschläge für einen griffigen Slogan auf dem Werbeplakat.

> >Entwirf ein Szenario für einen Open-Air-Kinoabend in einem Park. Unter den Gästen können auch Jugendliche ab etwa 14 Jahren sein. Welche Filme wären geeignet und wie sorgst du für das leibliche Wohl der Gäste?

> >Entwickle ein Konzept für eine nachhaltige Modenschau in einem alten Industriegebäude. Welche Materialien würdest du für die Bühne und Kulissen verwenden? Wie könnten Licht und Musik die Präsentation unterstützen?

3.5Reden und Rhetorik

> >Formuliere den Einstieg einer Rede für eine Feier zum Universitätsabschluss, der die Bedeutung von lebenslangem Lernen betont. Beginne mit einer lustigen kurzen Geschichte.

> >Schreibe einen Abschnitt einer Motivationsrede, die Menschen dazu inspiriert, ihre Träume zu verfolgen, egal wie alt sie sind und welches Einkommen sie haben.

> >Entwirf den Anfang einer geschichtlichen Rede, die an die ersten Schritte des Menschen auf dem Mond erinnert und die Bedeutung dieses Moments für die Menschheit hervorhebt.

> >Formuliere eine eloquente Danksagung für einen Mentor oder Lehrer, der einen tiefgreifenden Einfluss auf das Leben des Redners hatte.

> >Zeige mir anhand von Beispielen, welche rhetorischen Mittel für eine anspruchsvolle Rede vor gebildetem Publikum sinnvoll einzusetzen sind.

> >Gib mir eine Einleitung für eine Rede, die die Zuhörer davon überzeugt, die Bedeutung von Impfungen insbesondere gegen Grippe anzuerkennen.

> >Verfasse einen berührenden Abschnitt einer Eulogie für einen verstorbenen Freund, der stets als Quelle der Freude und Inspiration diente.

> >Entwickle einen Auszug aus einer Rede eines Stadtverwalters zur Krisenbewältigung nach einem großen Naturereignis, um Hoffnung und

Gemeinschaftssinn zu fördern.

> >Entwirf die ersten Zeilen einer Rede zur Eröffnung eines Charity-Events zur Unterstützung von Kindern in Bildungsnotlagen, in der du die Bedeutung von Zugang zur Bildung hervorhebst.

> >Gib mir eine strukturierte Übersicht für eine Antrittsrede eines neu gewählten Schulsprechers, der den Schülern mehr Mitspracherecht verspricht.

> >Skizziere den Abschluss einer Rede zur Eröffnung einer Kunstausstellung, die den Stellenwert der Kunst in Zeiten sozialer Unruhen betont.

> >Erstelle den Einstieg eines Fachvortrags über die Bedeutung von nachhaltiger Energie für zukünftige Generationen und den Planeten.

> >Wie erreicht man es, dass die eigene Rede satirisch wirkt?

> >Entwirf einen leidenschaftlichen Abschnitt, der die revolutionären Eigenschaften eines neu entwickelten Technologieprodukts für die Medizin beschreibt.

> >Integriere eine Anekdote in eine Rede über die Kraft und Bedeutung von Gemeinschaft und Zusammenarbeit in einem multikulturellen Umfeld.

> >Ein langjähriger Mitarbeiter hält eine Abschiedsrede: Formuliere einen herzlichen Abschnitt, in dem ein Mitarbeiter, der in den Ruhestand geht, seine Dankbarkeit und seine schönsten Erinnerungen an die Kollegen teilt.

> >Digitaler Datenschutz ist wichtig, gerade in Zeiten von KI-Sprachmodellen und Deep Learning: Skizziere das kraftvolle Ende einer Rede, die betont, wie entscheidend es ist, in der digitalen Ära persönliche Daten zu schützen und die Privatsphäre zu wahren.

3.6Marketing und Werbung

> >Erstelle zehn einzigartige Markennamen für ein umweltfreundliches Reinigungsprodukt.

> >Schreibe einen einnehmenden 50-Wort-Werbetext für ein neu eingeführtes Sportgetränk.

> >Beschreibe eine innovative Strategie zur Markenpositionierung für

ein nachhaltiges Modebabel, das den Massenmarkt anspricht.

> >Entwirf einen einnehmenden Social-Media-Post für den Start eines neuen Online-Buchclubs mit 200 Wörtern.

> >Formuliere einen auffälligen Mail-Betreff und einen kurzen E-Mail-Text, um Kunden über einen bevorstehenden Totalräumungsverkauf in einem Elektronikgeschäft zu informieren.

> >Entwickle fünf einprägsame Slogans für eine App, die Benutzern hilft, ihre Schlafgewohnheiten zu verbessern.

> >Skizziere die ersten fünf Schritte eines Marketingplans für ein neues veganes Restaurant in einer Großstadt. Zielgruppe sind sehr junge Leute mit geringer Kaufkraft.

> >Verfasse eine ansprechende Produktbeschreibung für ein innovatives, wiederverwendbares Wasserfiltersystem für Wanderer.

> >Beschreibe ein kreatives Konzept für eine Werbeanzeige, die die Vorteile von solarbetriebenen Ladeeinrichtungen für Smartphones hervorhebt.

> >Entwirf eine Struktur für eine Infografik, die die Vorteile von biologisch abbaubaren Verpackungen gegenüber herkömmlichen Kunststoffverpackungen darstellt.

> >Skizziere eine Strategie für eine Influencer-Kampagne, die ein neues, umweltfreundliches Sneaker-Brand promotet. Welche Art von Influencern würdest du ansprechen und warum?

> >Beschreibe eine innovative Guerilla-Marketing-Aktion für eine neue Fahrrad-Sharing-App in einer Großstadt, die Aufmerksamkeit erregt und virales Potenzial hat.

> >Gib mir eine Überschrift und einen kurzen Teaser für einen Blogartikel, der die Bedeutung von nachhaltigem Tourismus hervorhebt und dabei ein neues Öko-Hotel-Brand bewirbt.

> >Beschreibe das Storyboard für einen 30-sekündigen Video-Clip für eine App, die Menschen dabei hilft, ihre CO_2-Emissionen zu Hause und im Alltag zu reduzieren.

> >Entwickle eine Idee für ein Kundenbindungsprogramm für einen Online-Bio-Lebensmittelhändler. Wie könnten Kunden dazu motiviert werden, regelmäßig zu bestellen?

> >Verfasse einen 20-Sekunden-Spot für einen Podcast, der eine neue Meditations-App bewirbt. Wie könntest du die Hörer überzeugen, die

App auszuprobieren?

> >Plane eine Serie von fünf Instagram Stories für ein neues, alkoholfreies Craft-Bier. Welche Botschaften würdest du in jeder Story vermitteln?

3.7 Web-Programmierung mit HTML, CSS und JavaScript

> >Erstelle ein HTML- und CSS-Layout für eine Landing-Page, das sich automatisch an verschiedene Bildschirmgrößen anpasst, von Mobilgeräten bis zu Desktop-Computern, also im Responsive Design.

> >Verfasse einen einfachen JavaScript-Code, um ein Dropdown-Menü zu erstellen, das beim Klicken auf einen Button erscheint und verschwindet.

> >Wie könntest du mithilfe von CSS und JavaScript einen einfachen Parallax-Scrolling-Effekt für eine Webseite erstellen?

> >Zeige, wie man CSS-Transitions und -Animationen verwendet, um Hover-Effekte für Schaltflächen zu erstellen, die ihre Farbe und Größe ändern und erstelle einen Beispiel-Code.

> >Verfasse ein JavaScript-Script zur Validierung von Eingabefeldern eines Kontaktformulars, um sicherzustellen, dass alle Felder ausgefüllt sind und die E-Mail-Adresse gültig ist.

> >Verwende CSS Custom Properties (Variablen), um ein Farbschema für eine Webseite zu erstellen, das mit CSS-Variablen einfach geändert werden kann.

> >Ich brauche ein Modal-Fenster auf meiner Website: Erstelle einen JavaScript-Code, um ein Modal-Fenster anzuzeigen, das erscheint, wenn ein Benutzer auf eine Schaltfläche klickt, und sich schließt, wenn er außerhalb des Fensters klickt.

> >Verfasse ein CSS-Layout mit Flexbox für eine Galerieseite, die Bilder in verschiedenen Größen zeigt, sich aber an einheitlichen Abständen ausrichtet.

> >Ich möchte wissen, wie dynamisches Laden von Inhalten umgesetzt wird. Wie könntest du JavaScript verwenden, um Inhalte dynamisch von einem Array zu laden und sie in einer Liste anzuzeigen, sobald die Seite geladen ist?

> >Demonstriere, wie man SVG-Grafiken in HTML einbindet und sie

mithilfe von CSS-Stilen verändert.

> >Ich brauche einen Sticky Navigation Bar auf meiner Website. Erstelle einen HTML-Code für eine Navigationsleiste. Verwende anschließend CSS, um sie an der Oberseite des Bildschirms zu fixieren, und JavaScript, um ihre Farbe zu ändern, sobald der Benutzer über eine bestimmte Scroll-Position hinausscrollt.

> >Zeig mir, wie man eine Bildergalerie mit Lightbox-Effekt erzeugt. Wie könnte man eine Bildergalerie erstellen, bei der das angeklickte Bild in einem größeren Format (Lightbox) dargestellt wird und die Möglichkeit bietet, durch alle Bilder zu navigieren?

> >Entwickle einen interaktiven Tooltip, der erscheint, wenn der Benutzer mit der Maus über einen speziellen Text oder ein Icon fährt. Dieser Hover-Tooltip sollte detaillierte Informationen oder Hinweise anzeigen.

> >Implementiere einen 'Dark Mode'-Umschalter für eine Webseite, der das gesamte Farbschema der Seite mithilfe von CSS-Variablen und JavaScript ändert.

> >Setze Smooth Scrolling auf meiner Webseite um. Integriere eine Funktion, mit der der Benutzer beim Klicken auf Navigationslinks sanft zu verschiedenen Abschnitten der Webseite gescrollt wird.

> >Entwickle ein dynamisches Echtzeit-Suchfeld, das die Ergebnisse aus einer Liste von Einträgen filtert und sie dem Benutzer dynamisch anzeigt, während er tippt.

> >Gestalte ein responsives Hamburger-Menü für Mobilgeräte. Bei einem Klick auf das Menüsymbol soll ein Seitennavigationsmenü mithilfe von CSS-Transitions und JavaScript erscheinen.

> >Erstelle eine Zeitleiste von Ereignissen (Animation-Timeline) mit HTML und CSS. Nutze anschließend JavaScript, um Animationen auszulösen, wenn der Benutzer zu einem bestimmten Punkt auf der Seite scrollt.

> >Implementiere eine 'Lazy Loading'-Funktion für eine Bildergalerie, bei der Bilder erst geladen werden, wenn sie in den sichtbaren Bereich des Bildschirms scrollen, um die Ladezeit und Performance zu optimieren.

3.8Lebensmittel, Rezepte und Ernährungsplanung

> >Gesunde Frühstücksideen: Erstelle fünf schnelle und nährstoffreiche Frühstücksrezepte für Personen, die morgens wenig Zeit haben, aber energiegeladen in den Tag starten möchten.

> >Ich brauche einige vegetarische und vegane Alternativen: Schlage drei Hauptgerichte vor, die normalerweise Fleisch enthalten, und überarbeite sie zu vegetarischen oder veganen Versionen, ohne an Geschmack zu verlieren.

> >Entwirf einen 7-tägigen Ernährungsplan für einen Marathonläufer im Trainingsmodus, der auf eine ausgewogene Zufuhr von Proteinen, Kohlenhydraten und Fetten abzielt.

> >Ich koche für Allergiker und benötige leckere Rezepte: Präsentiere drei Dessert-Rezepte, die glutenfrei, laktosefrei und nussfrei sind, aber immer noch köstlich schmecken.

> >Bitte stelle mir Ideen für eine kulinarische Weltreise zusammen: Stelle ein Menü mit Vorspeise, Hauptgericht und Dessert zusammen, bei dem jedes Gericht aus einem anderen Land oder Kontinent stammt und die Vielfalt der Weltküche zeigt.

> >In meiner Familie müssen bekocht werden: ein Vegetarier, ein Veganer und eine Person mit Laktoseunverträglichkeit. Bitte stell mir ein außergewöhnliches Menü zusammen, das ich für ein gemeinsames Festessen zubereiten kann.

> >Nachhaltige Ernährung mit nachhaltigen Zutaten sind mir wichtig. Liste zehn Lebensmittel auf, die sowohl gesund als auch umweltfreundlich sind, und erkläre, warum sie gut für uns und den Planeten sind.

> >Entwirf drei sättigende Low-Carb-Hauptgerichte, die weniger als 20 Gramm Kohlenhydrate pro Portion enthalten und sich für eine Low-Carb-Diät eignen.

> >Vorkochen für die Arbeitswoche: Stelle einen 5-tägigen Mahlzeitenplan für Berufstätige zusammen, der einfache, im Voraus zubereitete Gerichte enthält, die in der Mittagspause genossen werden können.

> >Empfehle drei erfrischende Getränkerezepte mit und ohne Alkohol, die ideal sind, um sich an einem heißen Sommertag abzukühlen.

> >Ich brauche eine Auswahl gesunder Snacks für meine Kinder. Bitte erstelle eine Liste von zehn gesunden Snack-Ideen, die Kinder lieben werden und die in weniger als zehn Minuten zubereitet werden können.

> >Gib mir ein paar saisonale Kochideen für die Zeit nach dem Sommerurlaub. Für welche Gemüse- und Obstsorten ist der Herbst bekannt? Erstelle ein dreigängiges Menü, das diese saisonalen Zutaten nutzt.

> >Präsentiere fünf vegetarische Hauptgerichte, die mindestens 20 Gramm Protein pro Portion enthalten und ideal für Fitness-Begeisterte sind.

> >Entwickle einen Ernährungsplan für eine Woche, der speziell auf die Bedürfnisse von Senioren abgestimmt ist, mit besonderem Augenmerk auf Kalzium und Vitamin D.

> >Ich möchte für Diabetiker kochen. Stelle drei Dessert-Rezepte vor, die für Diabetiker geeignet sind, wobei der Zuckergehalt minimiert und durch gesündere Alternativen ersetzt wird.

> >Welche Weine oder alkoholfreien Getränke würdest du zu einem scharfen mexikanischen Abendessen empfehlen? Erkläre deine Auswahl.

> >Sind fermentierte Lebensmittel wirklich gesund? Erläutere die Vorteile fermentierter Lebensmittel für die Darmgesundheit und präsentiere ein einfaches Rezept für selbst gemachtes Sauerkraut.

3.9Tipps für Gesundheit und Fitness sowie Trainingspläne

> >Entwickle einen ausgewogenen Fitnessplan mit Trainingsroutinen, der auf die spezifischen Bedürfnisse von Jugendlichen, Erwachsenen bzw. Senioren zugeschnitten ist. Berücksichtige dabei altersspezifische körperliche Fähigkeiten und Bedürfnisse.

> >Erstelle eine Einführung in Yoga für Anfänger, beschreibe verschiedene Yogastile und erstelle eine einfache Sequenz mit Bildern und Beschreibungen der Grundhaltungen, die für Einsteiger geeignet ist.

> >Ich bin Hobbyläufer und möchte meine Lauftechnik optimieren. Schreibe eine Schritt-für-Schritt-Anleitung zur Überprüfung und Verbesserung der Lauftechnik, einschließlich Tipps zu Haltung, Schrittlänge und Atmung.

> >Ich brauche Ernährungsstrategien für Sportler mit Erklärungen, wie Makro- und Mikronährstoffe die sportliche Leistung beeinflussen. Erstelle einen sechswöchigen Ernährungsplan vor, der auf das Trainingsziel 10-Kilometer-Lauf abzielt.

> >Ich interessiere mich für die Psychologie des Sports. Diskutiere, wie mentale Stärke und Achtsamkeit die sportliche Leistung beeinflussen können und biete Techniken an, um diese Fähigkeiten zu fördern.

> >Widerstandstraining kann man ohne Gewichte nur mit dem eigenen Körpergewicht durchführen. Zeige eine Reihe von Körpergewichtsübungen, die zu Hause durchgeführt werden können, und betone ihre Vorteile und Anpassungsmöglichkeiten.

> >Ich brauche Hilfe zur richtigen Erholung nach dem Training. Biete Ratschläge und Techniken zur Muskelregeneration, zur Vermeidung von Übertraining und zur Maximierung der Vorteile eines Trainingsprogramms.

> >Entwerfe einen Trainingsplan zur Gewichtsreduktion, zur Verbesserung der Herzgesundheit oder zur Steigerung der Knochenstärke und erkläre, wie und warum diese Übungen ausgewählt wurden.

> >Ich möchte Intervalltraining für Fortgeschrittene durchführen. Erläutere das Konzept des High-Intensity-Intervalltrainings (HIIT) und stelle einen detaillierten Plan für eine 30-minütige Session vor, die verschiedene Intensitätsstufen und Pausen beinhaltet.

> >Skizziere Übungen, die diskret am Schreibtisch durchgeführt werden können.

> >Entwickle einen Trainingsplan für Outdoor-Aktivitäten im Winter.

> >Entwirf einen HIIT-Plan für zu Hause ohne zusätzliche Ausrüstung. Der Plan sollte vier Wochen umfassen und drei Trainingstage pro Woche beinhalten.

> >Zeige eine Reihe von Dehnübungen zur Steigerung der Flexibilität für Radfahrer, die man auch nach langen Ausfahrten machen kann.

> >Beschreibe kreative Übungen und Spiele, um Kinder zwischen drei und fünf Jahren zur Bewegung zu motivieren. Verpacke die Übungen in eine motivierende Geschichte.

> >Entwirf einen Ernährungsplan für den Muskelaufbau eines 55-jährigen Mannes, der zweimal in der Woche ins Fitnessstudio geht, sonst aber nur im Büro arbeitet.

> >Unterstreiche die Vorteile der Meditation im Sport und stelle einfache Übungsansätze vor, die besonders für Marathonläufer geeignet sind.

> >Wie würdest du einen 30-Tage-Trainingsplan erstellen, der sowohl

Krafttraining als auch Cardio beinhaltet, um einen ganzheitlichen Ansatz für Fitness zu bieten?

> >Ernährung und Energie sind untrennbar. Welche Superfoods und Ernährungstipps würdest du empfehlen, um den ganzen Tag über ein hohes Energielevel zu halten, ohne zu viel Carbs zu essen?

> >Ich brauche eine hohe mentale Stärke im Training. Welche Techniken oder Routinen könnten dabei helfen, mentale Ausdauer beim Sport zu entwickeln und die Motivation aufrechtzuerhalten?

> >Wie kann man mit Alltagsgegenständen effektive Work-outs für zu Hause gestalten, wenn man nicht ins Fitnessstudio gehen kann?

> >Ich muss dringend meine Balance verbessern. Welche Tipps würdest du mir geben, um einen gesunden Lebensstil und effektive Übungen in einen hektischen Alltag zu integrieren?

> >Ich möchte endlich mit dem Yoga beginnen. Skizziere eine einfache Yoga-Routine, die darauf abzielt, Stress abzubauen und vor allem meine Flexibilität zu erhöhen.

> >Kannst du mir helfen, mein Lauftraining zu optimieren? Welche Techniken und Ratschläge würdest du einem Anfänger geben, der sich auf seinen zweiten Halbmarathon vorbereiten möchte?

> >Wie hoch ist die Bedeutung der Regeneration nach dem Training? Welche Strategien unterstützen die Muskelregeneration? Was kann man tun, um Übertraining zu erkennen und zu vermeiden?

> >Wie hängen Hydration und Fitness zusammen? Wie wichtig ist Hydration für die sportliche Leistung und welche Tipps würdest du geben, um während des Trainings gut hydriert zu bleiben? Wie erkenne ich einen Wassermangel im Training?

> >Ist Hormonyoga noch für eine Frau über 50 zu empfehlen? Wenn ja, dann stelle mir eine passende Yoga-Übungsreihe zusammen. Wenn nicht, begründe dies bitte.

> >Ich benötige einen langfristigen Plan, der Bewegung, Ernährung und mentales Wohlbefinden kombiniert, um einen ganzheitlichen Ansatz für Gesundheit und Fitness zu bieten.

> >Wie würdest du einen 15-minütigen Morgen-Work-out-Plan gestalten, der den Kreislauf ankurbelt und einen perfekten Start in den Tag bietet?

> >Welche Snack-Ideen könntest du vorschlagen, die vor oder nach

dem Training ideal sind, um den Körper mit Energie zu versorgen und die Muskelregeneration zu unterstützen?

>>Ich will in das Intervalltraining einsteigen und dies zu Hause durchführen. Kannst du einen anspruchsvollen HIIT-Plan (High Intensity Interval Training) für zu Hause skizzieren, der ohne zusätzliche Ausrüstung auskommt?

>>Welche kreativen Übungen und Spiele könntest du Eltern empfehlen, um ihre 8- bis 10-jährigen Kinder spielerisch an Bewegung heranzuführen?

>>Ich möchte den Muskelaufbau mit gezielter Ernährung unterstützen. Welche Ernährungstipps und -pläne würdest du jemandem geben, der gezielt Muskelmasse aufbauen möchte?

3.10Originelle und kreative Geschenkideen für jeden Anlass

>>Wie könnte man eine persönliche Zeitkapsel gestalten, die der Beschenkte erst in zehn Jahren öffnen soll? Was könnte man darin verstauen, um die heutige Zeit zu konservieren?

>>Welche einzigartige virtuelle Realitätserfahrung könnte man für jemanden kreieren, der leidenschaftlich gern in ferne Länder reist, aber aktuell wegen einer schlimmen Krankheit nicht reisen kann?

>>Ich habe drei erwachsene Töchter zwischen 20 und 25, denen ich zu Weihnachten unterschiedliche, aber gleichwertige Geschenke machen möchte. Gib mir mindestens fünf Tipps, wie ich dieses Problem lösen kann.

>>Ich möchte einen Hobbykoch originell beschenken, aber nicht zu viel Geld für ein Geschenk ausgeben. Entwickle ein Konzept für ein einzigartiges DIY-Gewürzset. Welche exotischen Gewürze sollen enthalten sein und wie könnte man sie ansprechend präsentieren?

>>Für jemanden, der gern Neues lernt und Lernabenteuer liebt: Wie könnte man ein Paket aus verschiedenen Rätseln, Lehrbüchern und DIY-Kits zusammenstellen, um eine neue Fähigkeit wie das Jonglieren zu erlernen?

>>Ich möchte aus gemeinsamen Erinnerungen Kunst gestalten: Beschreibe, wie man aus gemeinsamen Fotos oder Briefen eine künstlerische Collage oder ein Skulptur-Objekt erstellen könnte.

> >Ich brauche kreative Vorschläge für ein musikalisches Geschenk. Wenn man für einen Musikliebhaber zum 40. Geburtstag eine Playlist erstellen könnte, die wie eine Reise durch verschiedene Jahrzehnte und Musikstile wirkt, welche Lieder wären enthalten?

> >Entwirf ein Konzept für ein ’blind date with a book’ als originelles Geschenk für einen leidenschaftlichen Buchliebhaber. Wir wollen dazu ein Buch als Geschenk verpacken, dieses aber nur anhand von Stichworten oder einem kurzen Rätsel beschrieben. Welches Buch würde das sein und wie würdest du es beschreiben?

> >Wie könnte man eine Karte eines Ortes als personalisierte Landkarte gestalten, der für den Beschenkten von besonderer Bedeutung ist, und dabei besondere Momente oder Erinnerungen an diesem Ort hervorheben?

> >Wie könnte man ein Geschichtenbuch der Erinnerungen gestalten, in dem man kleine Geschichten und Anekdoten aus dem Leben des Beschenkten zusammenträgt, vielleicht sogar in Form von kurzen Comics oder Illustrationen?

> >Entwirf ein DIY-Set für Pflanzenliebhaber, mit dem sie ihre eigene Mini-Oase oder ein kleines Kräutergärtchen zu Hause anlegen können. Welche Samen und Accessoires könnten enthalten sein?

> >Wir wollen ein ganz besonderes Retro-Erlebnis kreieren. Wie könnte man einen Retro-Abend gestalten, der die Beschenkten in die 70er-Jahre zurückversetzt? Was wäre in einem solchen Paket enthalten? Und wie könnte man die Umgebung passend gestalten?

> >Beschreibe ein Set von Rätseln oder Hinweisen als personalisierte Rätseljagd, die zu einem speziellen Ort oder einem Geschenk führen, das am Ende der Jagd wartet.

> >Welche spannenden Fakten, Rätsel oder Denkaufgaben könnten in einer ’Box des Wissens’ enthalten sein? Der Beschenkte ist 30 Jahre alt, Akademiker und leidenschaftlicher Angler.

> >Ich möchte ein Kunst- und Kreativset an einen Hobbykünstler verschenken. Welche Materialien und Anleitungen könnte man zusammenstellen, um einem Kunstliebhaber den Einstieg in eine neue Technik wie die Aquarellmalerei zu ermöglichen?

> >Hilf mir bei der Zusammenstellung eines Globetrotter-Paketes für einen jungen Menschen mit sehr großer Reiselust. Wie könnte man ein ’Reise-von-zu-Hause-aus’-Paket mit kulinarischen, musikalischen und kulturellen Überraschungen aus verschiedenen Ländern gestalten?

> >Mein Vater wird 70 Jahre alt, liebt den Nachthimmel über alles und soll ein Sternengucker-Paket zum Geburtstag erhalten. Wie könnte man ein Paket für Amateurastronomen zusammenstellen, inklusive Sternenkarte, Teleskop und Tipps für die besten Beobachtungsorte?

> >Ich denke an ein traditionelles Familienrezept oder ein besonderes Gericht. Wie könnte ich dieses in einem kreativen, kostengünstigen Geschenk für meine Ehepartnerin umsetzen, ohne dass ich dabei selbst koche?

> >Welches alte Hobby oder Kunsthandwerk könnten Sie wiederbeleben, um ein einzigartiges und bedeutsames Geschenk zu schaffen?

> >Welche Geschenke könnten Sie direkt aus der Natur oder aus Ihrem Garten kreieren, die sowohl schön als auch nützlich sind?

> >Ich möchte meinem Lebenspartner eine ganz persönliche musikalische Playlist schenken. Wie könnte ich so eine Playlist zusammenstellen, die die Geschichte unserer Beziehung erzählt?

> >Welche handgefertigten Beauty- oder Wellnessprodukte könnte ich selbst herstellen, um eine Frau mittleren Alters zu beschenken?

> >Ich verschenke gern Zeit statt Materielles. Welche Ideen gibt es dafür? Welche geteilten Erfahrungen oder gemeinsamen Momente kann ich verschenken? Und wie präsentiere ich das Geschenk in einem Gutschein?

> >Ich möchte ein Geschenk für einen passionierten Segler. Was kann ich selbst kreieren oder anbieten, um das Hobby zu unterstützen, ohne viel Geld dafür auszugeben?

> >Welche dekorativen Elemente könnten Sie mit einfachen Materialien für den Beschenkten basteln, um sein Zuhause zu verschönern?

> >Stelle Ideen für geschenkte Zeit und gemeinsame Aktivitäten vor.

> >Skizziere einen Plan für ein DIY-Geschenkset mit einem bestimmten Thema.

> >Entwickle Vorschläge für besonders nachhaltige Geschenkideen.

> >Präsentiere Konzepte und Bastelanleitungen für personalisierte Schmuckstücke oder modische Accessoires.

> >Stelle Ideen für Bücher oder Zeitschriften vor, die als besonderes Geschenk dienen können.

> >Entwirf eine Liste von Gadgets oder Technikspielzeugen für Technikliebhaber.

3.11Reiseziele, Reiseplanung und Reiseführer

> >Nenne mir fünf weniger bekannte, aber sehenswerte Orte in Europa und beschreibe kurz, was diese unentdeckten Perlen so besonders macht.

> >Erstelle eine 10-tägige Reiseroute durch Japan, die sowohl kulturelle als auch natürliche Attraktionen berücksichtigt.

> >Gib mir Tipps und Tricks, um beim Reisen mit dem Rucksack in Südamerika Geld zu sparen, ohne auf Qualität oder Erlebnisse zu verzichten.

> >Beschreibe ein Wochenendtrip-Konzept, bei dem es ausschließlich um das Erkunden von regionalen Köstlichkeiten in Niederösterreich geht.

> >Erstelle einen detaillierten Reiseführer für einen Tag in Rom, bei dem der Schwerpunkt auf dem antiken Rom liegt.

> >Empfehle mir eine Reiseroute für einen 2-wöchigen Abenteuerurlaub in Neuseeland, inklusive Outdoor-Aktivitäten und atemberaubenden Orten.

> >Erstelle einen Reisevorschlag für einen Karnevalliebhaber mit geringem Budget.

> >Liste fünf Strände auf, die oft übersehen werden, aber zu den schönsten der Welt gehören, und beschreibe, was sie so besonders macht.

> >Stelle einen Guide für einen nachhaltigen Urlaub in Costa Rica zusammen, der sowohl die Umwelt als auch die lokale Gemeinschaft berücksichtigt.

> >Gib einen Überblick über drei einzigartige kulturelle Festivals in Europa, die jährlich stattfinden, und was die dort Besucher erwartet.

> >Erkläre die besten Strategien und Ziele für Alleinreisende, die sowohl sicher als auch erfüllend sind, und betone dabei die Vorteile des Alleinreisens.

> >Skizziere einen 3-tägigen Reiseplan für Prag, wobei der Schwerpunkt auf historischen und architektonischen Wundern liegt.

> >Beschreibe eine 7-tägige Route durch die schottischen Highlands,

inklusive Wanderrouten, sehenswerten Burgen und besten Übernachtungsmöglichkeiten.

> >Erstelle eine detaillierte 2-wöchige Route für das Inselhüpfen in Griechenland, von bekannten zu weniger besuchten Inseln, mit Tipps für lokale Spezialitäten und versteckte Buchten.

> >Stelle mir ein einzigartiges Wüstenabenteuer zusammen. Skizziere dazu eine 10-tägige Abenteuerreise durch Marokkos Sahara, mit einem Mix aus Kameltrekking, historischen Kasbahs und lokalen Kulturerlebnissen.

> >Gib Empfehlungen für die Top-5-Tauchspots auf den Malediven, inklusive der besten Reisezeit und der zu erwartenden Meeresbewohner.

> >Ich will im Urlaub hauptsächlich architektonische Meisterwerke besuchen. Stelle einen Reiseplan für einen 4-tägigen Besuch in Barcelona zusammen, mit einem besonderen Fokus auf die Werke von Antoni Gaudí und modernistischen Bauwerken.

> >Meine Kinder sind acht und zehn, ich möchte trotzdem eine Abenteuerreise mit ihnen unternehmen. Nenne mir fünf Ziele mit passenden Aktivitäten.

> >Skizziere eine Winterreise nach Lappland, Finnland, mit dem Ziel, die Aurora borealis zu sehen. Betone dabei optimale Beobachtungsstandorte, Sami-Kultur und typische Winteraktivitäten.

3.12Finanzen, Investment und Vorsorge

> >Erläutere mir die Grundlagen der Investitionspsychologie. Welche Emotionen und kognitiven Verzerrungen können Investmententscheidungen negativ beeinflussen und welche Strategien können solche Fallstricke vermeiden?

> >Ich suche nach möglichen Szenarien für meinen Ruhestand. Zeichne zwei verschiedene finanzielle Wege in den Ruhestand: einen traditionellen Weg mit festem Einkommen und Renteneinzahlungen, und einen alternativen Weg durch selbstständige Investitionen und Passiveinkommen.

> >Bitte gib Tipps für die Finanzplanung junger Erwachsener. Erstelle einen Leitfaden für Personen in ihren 20ern, um den Wert des Sparens und Investierens früh im Leben zu verstehen, einschließlich Tipps zum Schuldenabbau und zur Altersvorsorge.

> >Diskutiere die Grundlagen von Kryptowährungen und Blockchain, ihre Rolle im modernen Finanzsystem und Überlegungen für potenzielle private Investoren, die überlegen, in diesen Bereich einzusteigen.

> >Wie sieht ein gutes allgemeines Risikomanagement für Investition aus? Erläutere, wie Risikotoleranz bewertet wird und wie sie die Auswahl von Investitionen beeinflusst. Stelle Strategien vor, um das Risiko in einem Portfolio effektiv zu managen.

> >Meinen Vermögensaufbau möchte ich über Immobilien realisieren. Untersuche die Vor- und Nachteile des Immobilieninvestments im Vergleich zu anderen Anlageformen und skizziere einen Plan für den Kauf des ersten Mietobjekts.

> >Ich möchte bei meinen Finanzen die Ethik immer fest im Blick behalten. Diskutiere die Bedeutung ethischer Investitionen und wie man Unternehmen und Fonds identifiziert, die in Übereinstimmung mit bestimmten ethischen oder nachhaltigen Prinzipien arbeiten.

> >Untersuche, wie globale wirtschaftliche Trends und geopolitische Ereignisse die Investmentlandschaft beeinflussen können und wie Anleger auf solche Veränderungen reagieren sollten.

> >Gib mir bitte einen Überblick über verschiedene Fondsarten. Zeige die Unterschiede und Charakteristika von ETFs, Mutual Funds und Indexfonds auf und diskutiere die Vor- und Nachteile jeder Anlageform.

> >Ich möchte über steuerliche Aspekte unterschiedlicher Formen der Kapitalanlage informiert werden. Erläutere, wie verschiedene Anlageformen besteuert werden und welche Strategien Anleger nutzen können, um ihre steuerliche Belastung in Deutschland zu minimieren.

> >Die Pensionsplanung befindet sich im permanenten Wandel. Untersuche, wie sich die Landschaft der Altersvorsorge im Laufe der Jahre verändert hat, von traditionellen Rentenplänen bis hin zu modernen 401(k)s und IRAs.

> >Die Zentralbanken spielen im Investmentumfeld offenbar eine wichtige Rolle. Diskutiere, wie Zentralbankentscheidungen, Zinssätze und geldpolitische Maßnahmen die globalen Märkte beeinflussen.

> >Kann ich auch Rohstoffe für Investitionen nutzen? Erläutere die Grundlagen des Investierens in Rohstoffe, einschließlich Gold, Öl und seltene Erden, und ihre Rolle in einem diversifizierten Portfolio.

> >Automatisiertes Investieren und Robo-Advisors klingen nach einer spannenden Möglichkeit für Anleger ohne Fachwissen. Wie funktionieren

diese automatisierten Anlageplattformen und für wen sind sie am besten geeignet?

> >Ich will für den Pflegefall vorsorgen. Welche wirtschaftlichen Herausforderungen im Zusammenhang mit Langzeitpflege gibt es und welche Strategien zur finanziellen Absicherung kannst du empfehlen?

4ChatGPT vs. Google: Ein Vergleich

Äpfel darf man niemals mit Birnen vergleichen, das weiß nun wirklich jedes Kind, und insofern ist die Überschrift »ChatGPT vs. Google« eigentlich unsinnig. Der Grund, warum wir diesem Thema trotzdem ein ganzes Kapitel widmen, ist, dass beide Dienste bei den Nutzern in einer gewissen Konkurrenzsituation stehen. Tatsächlich ringt ChatGPT dem Monopolisten Google immer mehr Marktanteile ab, spätestens seit der Chatbot in die Bing-Suche integriert wurde. Damit tritt praktisch etwas in eine direkte Konkurrenz, was theoretisch eigentlich gar nicht in einem Wettbewerb stehen dürfte, weil es sich um zwei grundverschiedene Anwendungen bzw. Dienste mit grundlegend unterschiedlicher Arbeitsweise handelt. Wer ChatGPT als Ersatz für Google betrachtet, hat die Suchmaschine und ihren eigentlichen Wert noch nicht richtig verstanden. Andererseits gibt es tatsächlich einige Anfragen, für die Google bisher die erste und einzige Anlaufstelle war, die sich künftig mit ChatGPT einfacher und smarter erledigen lassen. Dennoch ist der Chatbot kein Ersatz für eine Suchmaschine, wird es auch niemals werden und wir Anwender müssen lernen, welches Werkzeug sich für eine vorliegende Aufgabe am besten eignet.

Warum ist Google das Beispiel, wenn es doch so viele andere Suchmaschinen gibt? Weil Google weltweit die weitaus größten Marktanteile hat, die Internetsuche grundlegend revolutioniert hat und nach Einschätzung vieler Branchenkenner technologisch weit voraus ist. Stand heute verwenden viele »Suchmaschinen« die Google-Suche oder den Google-Index im Hintergrund, Google-Technologien finden sich mit mehr oder weniger zeitlicher Verzögerung in anderen Suchdiensten wieder, Google-Werbeformate werden von anderen Anbietern nachgeahmt.

Wir wollen in den folgenden Abschnitten kurz erklären, was eine Suchmaschine ausmacht und wie sie grundlegend arbeitet. Anschließend betrachten wir die Entwicklung bei Google im Schnelldurchlauf und werfen einen Blick auf wichtige Updates und Technologien. Ziel ist es zu zeigen, mit welch ausgeklügelten und über Jahre gereiften Methoden die Google-Suche arbeitet. Am Ende dieses Abschnittes haben Sie eine Vorstellung davon, welche hohe Qualität in den Google-Suchergebnissen steckt und wie verlässlich die Quellen auf den obersten Rängen sind.

Mit diesem Wissen ausgestattet können wir uns dann den wichtigsten Unterschieden zwischen den Antworten eines Sprachmodells und der Ergebnisliste einer Suchmaschine widmen. Tipps, für welche Anfragen

Sie künftig welches Werkzeug nutzen, erübrigen sich dann: Mit dem Wissen über die Zusammenhänge wählen Sie für jede Frage künftig automatisch den richtigen Dienst.

4.1Wie funktioniert eine moderne Suchmaschine?

Indexbasierte Suchmaschinen wie Google sind komplexe Softwareprogramme, die selbstständig das Internet durchsuchen und daraus ihren Datenbestand automatisch erzeugen, erweitern und ständig aktualisieren, um Nutzern dann relevante Ergebnisse auf beliebige Suchanfragen liefern zu können. Solche indexbasierten Suchmaschinen bestehen im Wesentlichen aus drei Komponenten, wobei nur die dritte Komponente bei den Nutzern überhaupt in Erscheinung tritt.

1. Der erste Teil bzw. der erste Schritt ist die **Datengewinnung** mit dem **Webcrawler**, auch einfach Crawler, Robot oder Spider genannt. Aufgabe des Webcrawlers ist die Erweiterung des Dokumentenindex (eine sehr lange Liste aller URLs) durch die laufende Erschließung unbekannter Ressourcen. Weitere Aufgabe des Crawlers ist die Pflege des Index, weil dort enthaltene Dokumente laufend auf Änderungen geprüft und aktualisiert werden müssen.
2. Nun folgt die **Datenanalyse**, **Datenaufbereitung** und **Datenverwaltung** als logischer Schritt nach dem Crawling. Die gefundenen Dokumente werden geprüft, analysiert und normalisiert (von HTML in ein suchmaschineninternes Metaformat umgewandelt). Nach diversen Prüfungen, die jede URL vor der Weiterverarbeitung durchlaufen und bestehen muss, erfolgt die Speicherung der Seite und der Aufbau einer durchsuchbaren Datenstruktur für die spätere Abfrage (*Information Retrieval*).
3. Die **Beantwortung** von konkreten **Suchanfragen** erfolgt im **Query Prozessor**, dem dritten und letzten Modul, mit dem eine Suchmaschine für ihre Nutzer eigentlich erst in Erscheinung tritt. Suchanfragen werden analysiert, eine Selektion der Treffer wird durchgeführt und dann das Ranking für die Suchergebnisseite bestimmt. Hier wird auch die Werbung eingespielt.

In den Urzeiten des World Wide Webs waren Suchmaschinen teilweise noch Kataloge bzw. thematisch sortierte Verzeichnisse, in die man neue Domains händisch eintragen musste, um überhaupt eine Listung zu bekommen. Yahoo ist ursprünglich als Webkatalog gestartet, das exponentielle Wachstum des WWW hat diese Technologie aber schnell überholt. Moderne Suchmaschinen sind allesamt indexbasiert, bauen

ihren Datenbestand selbst auf und aktualisieren ihn laufend. Indexbasierte Suchmaschinen wie Google arbeiten schnell, die Inhalte sind in der Regel ziemlich vollständig und gemessen an der Menge der Informationen auch sehr aktuell. Fehler und Verzögerungen lassen sich naturgemäß nicht vermeiden, sind aber die Ausnahme. Der Dokumentenindex von Google enthält alle bekannten URLs, ist eine der wichtigsten Ressourcen der Suchmaschine und wird durch neue Links (oder Anmeldungen) ständig erweitert.

An dieser Stelle nochmals zurück zum Crawling. Die Datengewinnung mit dem Webcrawler folgt einem sehr einfachen Prinzip: Dieser startet auf einer beliebigen Seite und speichert alle dort gefunden Links, die zu weiteren Seiten innerhalb der Domain (interne Links) und zu weiteren URLs auf anderen Domains führen (externe Links). Alle Links werden mit dem Dokumentenindex abgeglichen, um herauszufinden, welche neu sind und welche aktualisiert werden müssen. Nach diesem einfachen Prinzip findet der Crawler einer Suchmaschine systematisch neue Seiten und neue Domains, wenn auch nur ein einziger Link darauf zeigt. Für einen umfassenden und aktuellen Index muss der Crawler sehr fleißig sein bzw. die Suchmaschine muss permanent eine gewaltige Rechenleistung bereitstellen und über sehr effektive Methoden und Algorithmen verfügen.

Nach der Datengewinnung mit dem Webcrawler beginnt die Aufbereitung der gefundenen Inhalte, zunächst aber muss jedes Dokument eine ganze Reihe verschiedener Prüfungen durchlaufen. Eine davon ist die Dublettenerkennung, es wird also geprüft, ob dieselben Inhalte oder wesentliche Teile davon schon im Index vorhanden sind. Das wäre für Google sogenannter *Duplicate Content* und der wird aus Platzgründen nicht nochmals gespeichert. Weitere Prüfungen suchen nach bestimmten Begriffen, die auf einer Blacklist stehen und die Google auch nicht im Index haben möchte. Hat eine URL diese und alle weiteren Prüfungen erfolgreich bestanden, wird sie als lokale Kopie im sogenannten *Repository* gespeichert. Das ist allerdings wieder nur ein Zwischenschritt, denn URLs im Repository sind wie die Seiten eines Buches schwer durchsuchbar und eine Google-Suchanfrage würde Stunden oder Tage beanspruchen und viel Rechenleistung binden.

Erst die verschiedenen Indizes beschleunigen den Zugriff bei der Suche bis in den Bereich von Millisekunden und geben Hinweise für das spätere Ranking. Dazu startet nun ein aufwendiger Prozess der Datenanalyse und Datenaufbereitung. Zuerst werden die vom Crawler bereitgestellten Seiten in ein gemeinsames Meta-Format konvertiert (Normalisierung), dann beginnt die eigentliche Datenanalyse (Parsing). Nach weiteren Schritten wie der Sprachidentifikation und der Wortidentifikation werden im Word-Stemming alle Wörter auf ihre Grundformen reduziert

und mit einem Wörterbuch abgeglichen. Erst im letzten Schritt wird eine durchsuchbare Struktur erzeugt, bestehend aus einer *Hitlist* (wo und wie wurde ein Keyword gefunden), einem *direkten Index* (einer Liste aller Keywords im untersuchten Dokument) und einem *indirekten Index*, der für jedes Keyword Verweise auf die Dokumente enthält, in denen es vorkommt. Der indirekte Index von Google entspricht grob gesprochen dem Index eines Buches mit Seiten-Verweisen auf die Fundstellen von Keywords. Damit ist auch der zweite sehr aufwendige Schritt abgeschlossen und die neue (oder aktualisierte Seite) wurde vollständig indexiert.

Die Verarbeitung von Suchanfragen und deren Beantwortung im sogenannten *Query Prozessor* ist der letzte Schritt, bei dem die Suchmaschine jetzt erstmals in Erscheinung tritt. Wir tippen (oder sprechen) unsere Anfrage in den Suchschlitz, die Suchmaschine nimmt die Suchanfrage entgegen und analysiert diese. Dabei laufen teilweise dieselben Prozesse ab, die wir beim zweiten Schritt der Datenaufbereitung schon kennengelernt haben, unter anderem muss die Sprache erkannt werden, in der die Frage gestellt wurde. Schließlich werden die für eine Antwort infrage kommenden Dokumente ausgewählt und die Reihenfolge der Ausgabe wird festgelegt. Dieser letzte Schritt – das Ranking der Suchergebnisse – gehört zu den magischen Rezepten einer Suchmaschine und wird von den Betreibern streng geheim gehalten. Je mehr über die Rankingfaktoren und deren Gewichtung bekannt ist, umso einfacher und stärker können diese manipuliert werden. Die Geschichte von Google ist daher auch eine Geschichte der (nie endenden und oft vereitelten) Manipulation oder Manipulationsversuche.

4.2Google – vom Start-up zum Monopolisten

Google ist in vielen Teilen der Welt zum Synonym für die Online-Suche und zum ersten Zugangspunkt für zahlreiche Online-Aktivitäten geworden. Heute ist die Suchmaschine der wichtigste Teil des Alphabet-Konzerns und neben YouTube die wichtigste Einnahmequelle des weltweit agierenden Unternehmens. Die Anfänge von Google reichen bis 1995 zurück. Ausgangspunkt war die Doktorarbeit von Larry Page an der Stanford University. Schon im Jahr 2000 war Google als Suchmaschine in 15 Sprachen nutzbar und verfügte über den weltweit größten Suchindex mit einer Milliarde URLs. Das Werbenetzwerk AdWords sicherte der Suchmaschine eine solide finanzielle Basis, bis heute lebt der gesamte Konzern hauptsächlich von den Werbeeinnahmen seiner

zahlreichen Dienste. Die Google-Suche wurde schnell ausgeweitet. 2005 folgten Google Maps und Google Analytics, 2008 der Google-Chrome-Browser mit der Google-Suche als Standard über die zentrale Suchleiste. Android und das Smartphone begannen einen beispiellosen Siegeszug und 2012 wurde Google Play als zentraler Ort zum Download von Android-Apps eingeführt.

Heute ist die Google-Suche schon lange keine einfache Websuche mit zehn Ergebnissen mehr, sondern hat sich zu einem umfangreichen, universellen Suchwerkzeug entwickelt, das auf verschiedenen Geräten für spezifische Suchen nach Dokumenten oder anderen Inhalten eingesetzt wird. Unterschiedliche Inhaltstypen wie Bilder, Videos und Nachrichten haben keine separaten Ergebnisseiten bekommen, sondern wurden als *Universal Search* direkt in die Haupt-Suchergebnisseite integriert. Zur Universal Search gehören heute neben Webseiten als traditionellen Suchergebnissen auch Bilder und Videos (insbesondere von YouTube), Karten und lokale Ergebnisse, News, Bücher und weitere Inhalte. Für produktbezogene Suchanfragen zeigt Google eine Leiste mit Shopping-Ergebnissen oder Produktlistenanzeigen (PLAs) an. Weil die Universal Search über viele Jahre schrittweise ausgerollt wurde, ist die Vielfalt der Google-Suchergebnisse den wenigsten Nutzern bewusst. Wir haben uns langsam und schrittweise daran gewöhnt. In den Google SERPs (Search Engine Result Pages, also den Google-Suchergebnisseiten) ist schlichtweg alles zu finden.

4.3Googles Wikipedia: Der Knowledge Graph

Der Google Knowledge Graph ist ein Teil der Universal & Extended Search und damit eine Anreicherung der organischen Google-Suchergebnisse. Der Knowledge Graph ist eine Wissensdatenbank, aus der Google relevante Informationen in einer Infobox neben seinen Suchergebnissen bereitstellt und dem Benutzer direkt antwortet.

Mit dem Google Knowledge Graph hat Google erstmals das Prinzip verlassen, nur auf Quellen zu verweisen, und baut eine eigene Wissensbasis mit eigenen Inhalten auf. Der Knowledge Graph erfasst Fakten über Entitäten (Personen, Orte, Organisationen, Dinge usw.) und deren Beziehungen zueinander, um Benutzern schnelle und präzise Antworten auf bestimmte Anfragen zu bieten wie zum Beispiel:

- Wie hoch ist der Mount Everest?
- Was ist die Hauptstadt von Venezuela?
- Wie heißt die Ehefrau von König Charles?

Wissens-Boxen und Wissens-Panels sind Formen der Darstellung des Knowledge Graphs, die an prominenter Stelle in den Google-Suchergebnissen erscheinen und die typischen Suchergebnisse um eigene Google-Inhalte erweitern. Natürlich stammen die Informationen im Panel stets aus vertrauenswürdigen Quellen, die Google kennt und zu denen es freien Zugang hat. Wikipedia ist wohl die wichtigste Anlaufstelle, aber auch andere offizielle Websites, lizenzierte Datenbanken und weitere vertrauenswürdige Online-Ressourcen. Wir sehen in folgenden Abschnitten noch, dass Google sehr großen Wert auf korrekte Informationen und vertrauenswürdige Quellen legt.

4.4Die Tricks von Google

Wie konnte Google in so kurzer Zeit zum führenden Suchdienst aufsteigen? Das Revolutionäre am neuen Suchdienst war der PageRank-Algorithmus, das ursprüngliche Thema der Doktorarbeit von Larry Page. Statt nur die Inhalte und Signale innerhalb einer Seite bzw. URL zu betrachten, analysiert und gewichtet dieser Algorithmus die externe Verlinkungsstruktur einer Seite. Auch hier ist das Grundprinzip wieder sehr einfach: Je mehr hochwertige Links auf eine Webseite verweisen, desto höher ist der PageRank dieser Seite und desto besser das Ranking bei Google. Wie in einer demokratischen Wahl wird jeder externe Link (Backlink) als Stimme gewertet und einem Kandidaten (der verlinkten URL) zugerechnet. Damit beginnt das Wettrennen um die meisten und besten Links, das noch immer andauert und wohl nie zu Ende sein wird.

Weil man mit einem guten Google-Ranking schon bald richtig gutes Geld verdienen konnte, nutzen viele kreative *Suchmaschinenoptimierer* dieses Prinzip dafür aus, massenhaft künstliche Backlinks zu produzieren und damit das Google-Ranking in ihrem Sinne zu beeinflussen. Die gut gemeinte Idee des PageRanks wurde massiv unterlaufen und die Qualität der Suchergebnisse schlechter. Google hat im Laufe der Jahre viele Updates für seinen Suchalgorithmus eingeführt, um die Qualität der Suchergebnisse kontinuierlich zu verbessern und Manipulationsversuche zu bekämpfen. Hier sind einige der wichtigsten Google-Updates und eine kurze Erklärung ihrer Arbeitsweise:

Panda (2011) war ein Filter in der Google-Suche, um Websites mit dünnem, dupliziertem oder geringwertigem Inhalt zu erkennen und herabzustufen. Panda war darauf ausgerichtet, Websites von hoher Qualität in den Suchergebnissen hervorzuheben und solche von geringerer Qualität zu reduzieren.

Penguin (2012) sollte Websites mit unnatürlichen Backlinks oder

übermäßig optimierten Anchor-Texten (der unterstrichene Teil eines textbasierten Links) erkennen und herabstufen. Penguin belohnte Websites mit natürlichem Linkaufbau und bestrafte solche mit offensichtlich gekauften und manipulierten Links.

Weitere Updates sorgten für eine höhere Qualität der Suchergebnisse und verfolgten dazu unterschiedliche Ansätze. Einige zielten auf ein besseres Verständnis von Suchanfragen, insbesondere bei komplexeren und gesprochenen Eingaben:

Hummingbird (2013) berücksichtigte den gesamten Kontext einer Suchanfrage, anstatt sich nur auf einzelne Keywords zu konzentrieren. Es nutzte erstmals die *semantische Suche*, um den Kontext und die Absicht des Nutzers besser zu verstehen.

BERT (2019) sollte Nuancen in den Eingaben und den Kontext der Suchanfragen noch besser verstehen und nutzte dazu erstmals ein Deep-Learning-Modell namens BERT (Bidirectional Encoder Representations from Transformers). Es gelang damit tatsächlich, den Kontext einer Suchanfrage, basierend auf allen anderen Wörtern der Eingabe, besser zu verstehen.

Eine ganz besondere Bedeutung nimmt das folgende Update ein. Es unterstreicht Googles vielfältigen und systematischen Ansatz, die positive Benutzererfahrung in den Mittelpunkt zu stellen. Als Schlüssel zu diesem Ziel wurden hochwertige Inhalte auf Websites ausgemacht, die nur von ganz bestimmten Personen oder Gruppen erstellt werden können. Wenn es um hochwertigen Content geht, wird Google richtig pedantisch und sieht sehr genau hin!

Das **E-A-T-Update** (2018) hat nichts mit kulinarischen Genüssen zu tun, sondern stellt erstmals die Begriffe *Expertise*, *Autorität* und *Vertrauen* (engl. *trust*) in den Vordergrund. Wer hat die Inhalte eines Artikels oder der gesamten Website verfasst? Hatte er oder sie das dafür notwendige Fachwissen? Ist der oder die Verfasserin eine wie auch immer anerkannte Autorität auf dem jeweiligen Gebiet? Und wie vertrauenswürdig sind die Inhalte und Informationen? Von diesem Update besonders betroffen waren sogenannte YMYL-Inhalte und Websites. Das Akronym steht für *Your money your life*, also für die Themen *Finanzen* und *Gesundheit*, weswegen es auch unter dem Begriff **Medic-Update** zu finden ist. Bei diesen Themen können falsche Informationen sehr negative Auswirkungen auf die Suchenden haben und deswegen wollte Google bei der Auswahl der Quellen sehr behutsam und umsichtig sein.

Mittlerweile wurde E-A-T noch um den Faktor *Experience* zu E-E-A-T ausgeweitet und wird jetzt sogar für die Beurteilung aller Websites herangezogen, also unabhängig vom Thema. Google möchte damit sicherstellen, dass nur verlässliche Informationen oben in den

Suchergebnissen angezeigt werden und die Nutzererfahrung immer sehr hoch ist. Deshalb hat die Suchmaschine diese und weitere Mechanismen entwickelt, um die selbst gesetzten Qualitätsziele immer besser zu erreichen. Auch Google beschäftigt eine große Anzahl an Clickworkern, die einzelne Websites besuchen und nach einem eigens dafür entwickelten Qualitätshandbuch (Search Quality Rater Guidelines) bewerten, ob und in welchem Maße die E-E-A-T Kriterien erfüllt sind.

4.5Arten von Suchanfragen

Hinter jeder Suchanfrage steckt eine Suchintention: Je nach dem erwarteten Ergebnis formulieren die Anwender ihre Suchanfrage in etwas anderer Form. Suchmaschinen haben selbst ein großes Interesse daran, immer den am besten passenden Content auszuspielen, weil das ihre Reputation steigert und langfristig Werbeeinnahmen generiert. Um die Intention der Nutzer zu verstehen, werden nicht nur die Keyword-Kombinationen der Sucheingabe ausgewertet, sondern auch alle verfügbaren Signale der Suche (z. B. die Suchhistorie) herangezogen.

Nur wenn eine Suchmaschine die Absicht hinter einer Anfrage erkennt, kann sie das passende Ergebnis ausspielen: Nutzer, die etwas kaufen möchten, bekommen Google-Shopping-Anzeigen und Links zu Marktplätzen oder Onlineshops mit einem entsprechenden Angebot. User, die nach Informationen suchen, finden eher einen Blogartikel oder ein Tutorial zum gewünschten Thema. Eine in der SEO-Szene gängige Typologie unterscheidet drei Arten von Suchanfragen: *transaktionsgetriebene* (transactional), *navigationsgetriebene* (navigational) und *informationsgetriebene* (informational) Suchen.

Transaktionale Suchanfragen haben eine klare Kaufabsicht oder lassen eine andere geplante Transaktion erkennen. Suchende wollen ein Produkt erwerben, egal ob kostenlos oder gegen Bezahlung. Andere wollen vielleicht eine Reise buchen oder eine Veranstaltung besuchen. Sie nutzen die Suchmaschine, um ein bestimmtes Produkt zu finden, auszuwählen, aufzurufen und es schließlich zu erwerben. Transaktionale Suchanfragen erkennt man an bestimmten Verben wie *kaufen*, *bestellen*, *buchen* oder *herunterladen*, aber auch daran, dass spezifische Produkte und Produktnamen in einem eindeutigen Kontext verwendet werden. Manche Suchanfragen sind zwar nicht ganz eindeutig, lassen jedoch eine Transaktionsabsicht zumindest vermuten. Die Suche nach einem Flug von A nach B, nach einem Konzert oder nach einem Hotel einer bestimmten Region deuten darauf hin, dass ein Suchender den Flug oder das Hotel auch tatsächlich buchen würde. Eine Suchmaschine kann

Preise angeben, Termine nennen oder Zeiträume bereitstellen.

Navigationale Suchanfragen, der zweite Typ, sind Suchen nach einer Website oder einer Quelle als direktes Ziel. Der Nutzer möchte ein bestimmtes Angebot aufrufen und verwendet die Suchmaschine als Hilfsmittel zur Navigation, weil die korrekte Schreibweise des Anbieters oder die Webadresse des Unternehmens nicht bekannt sind.

Der letzte Typ, die **informationalen Suchanfragen**, sind die ursprünglichste Form aller Suchanfragen. Schließlich war es die Grundidee der Suchmaschinen, Informationen gezielt in der großen Datenmenge des World Wide Webs zu finden und auf die Fundstellen zu verweisen. Informationsorientierte Suchanfragen sind an den Keyword-Kombinationen kaum noch zu erkennen, so vielfältig sind die Ansätze der Suchenden. Selbst typische W-Fragen sind kein sicheres Indiz dafür.

Die wichtige Erkenntnis daraus: ChatGPT kann weder transaktionale noch navigationale Suchanfragen beantworten, diese Typen bleiben immer den Suchmaschinen und ihrer besonderen Funktionsweise vorbehalten. Schon aus diesem Grund ist es für einen Abgesang der Suchdienste noch viel zu früh. Noch ein anderer Grund spricht dafür, dass die Suchmaschinen eine Zukunft haben: Transaktion und Navigation sind genau die beiden Typen, mit denen bei Google Ads und Google Shopping das meiste Geld für Alphabet verdient wird. Informationale Suchanfragen mögen vielleicht der Kernbereich der Suchmaschinen sein, sind aber deutlich schwieriger zu monetarisieren. Langfristig sind die Refinanzierungsoptionen ein echter Vorteil der Suchmaschinen und gleichzeitig der eingebaute Nachteil von Sprachmodellen. Überlegungen dazu gibt es bereits, aber noch weiß man nicht so recht, wo man bei ChatGPT & Co die Werbung am besten unterbringen kann.

Obwohl die drei genannten Arten von Suchanfragen längst zum allgemein akzeptierten Standard in der SEO-Szene gehören, möchte ich für den Vergleich mit ChatGPT eine etwas andere Typologie mit einfacheren Bezeichnungen verwenden. Natürlich nutzen wir Suchmaschinen immer, um Information zu bekommen, aber die Qualität dieser Infos ist sehr unterschiedlich. In der Praxis nutzen wir Google wohl hauptsächlich für diese Dinge:

1. **Single Information**
   Das sind einfache Informationen, die man meistens mit einem oder wenigen Worten bzw. mit einer Zahl beantworten kann. Wir hatten solche Fragen bereits:

   - Wie hoch ist der Mount Everest?
   - Was ist die Hauptstadt von Venezuela?
   - Wie heißt die Ehefrau von König Charles?

   Wie wir gesehen haben, hat Google dafür eine eigene Wissensbasis

und antwortet im Knowledge Graph. Schlüsselkonzepte sind Entitäten (Personen, Orte, Objekte), Attribute (Eigenschaft oder Information der Entitäten) und Beziehungen zwischen den Entitäten.

2.Complex Information
Hier geht es um Erklärungen und Definitionen, Konzepte und Zusammenhänge. Solche Antworten sind schlichtweg zu komplex für einen Knowledge Graphen. Wir finden die Antworten meistens auf spezialisierten Websites:

- Wie funktioniert ein Dynamo?
- Was versteht man unter Cancel Culture?
- Welche Ursachen führten zum Ausbruch des Ersten Weltkrieges?

Komplexe Fragen erfordern komplexe Antworten und Erklärungen, wie wir sie oft in Blogartikeln, Beiträgen oder Essays finden. Google antwortet nicht mehr selbst, sondern verweist auf die besten Spezialisten und anerkannten Quellen des jeweiligen Fachgebietes.

3.News
Auch hier geht es um Informationen, einfache oder komplexe, das gemeinsame Kennzeichen bei den News ist deren Aktualität. Beim Wetter sind es Stunden, beim Sport sind es Tage und in der Zeitgeschichte können Ereignisse im Zeitraum von Wochen immer noch im Rahmen der Aktualität sein.

4.Shopping
Das deckt sich natürlich weitgehend mit den vorher genannten transaktionalen Suchanfragen, ist von der Bezeichnung her aber vielleicht etwas griffiger. Konzertkarten, Flüge und Hotels fallen natürlich auch unter diese Kategorie.

5.Navigation
Die Navigation ist ein Sonderfall, und es könnte für unsere Fragestellung helfen, diesen Begriff wirklich sehr weit zu fassen: Mit dem Auto zum Ort X, auf dem Amt in das Zimmer mit den Anträgen für Y, mit dem Browser zur Website des Anbieters Z. Die Suchmaschine kennt immer den Weg! Klassische navigationale Suchanfragen wären damit natürlich auch abgedeckt.

Was davon kann ChatGPT oder ein anderes Sprachmodell? Natürlich die Bereiche *Single Information* und *Complex Information*, mehr aber auch nicht. News, Shopping und Navigation sind aus verschiedenen Gründen nicht möglich und auch in kommenden Versionen eines Sprachmodells nicht sinnvoll. Das kann eine Suchmaschine besser und sie wird es dank ihrer Arbeitsweise immer besser und verlässlicher können. Bleiben 1)

und 2) als informationale Suchanfragen.

Bei der *Single Information* herrscht eine gewisse Patt-Situation zwischen dem Google Knowledge Graphen und den Antworten von ChatGPT, beide antworten schnell und die Information ist leicht zugänglich. Weil ich die Antwort des Sprachmodells sicherheitshalber noch überprüfen müsste (Stichwort *Faktencheck*), kann ich mich aber auch gleich an die Suchmaschine wenden. Dort warten Bilder und Illustrationen, vielleicht auch Tabellen und Grafiken mit weiterführenden Informationen, und in jedem Fall erhalte ich jede Menge Links zur Vertiefung. Die Ehefrau von König Charles? Ich sehe Bilder von zwei Frauen und bin mit dem nächsten Klick mitten in der jeweiligen Geschichte. Also kein Vorteil für ein Sprachmodell, auch wenn es auf den ersten Blick vielleicht bequemer aussieht.

Das Blatt wendet sich bei 2), den komplexen Informationen, wenn es also um vielschichtige Erklärungen, anspruchsvolle Beschreibungen und tiefgehende Deutungen geht. Die Suchmaschine verweist mich naturgemäß auf statische Inhalte, die ich verstehen kann oder eben auch nicht verstehe. Bei komplexen Konzepten und Begriffen hat es manchmal sogar den Anschein, die Autoren der Artikel auf der ersten Google-Seite hätten ein gegebenes Problem auch nicht recht verstanden und nur voneinander abgeschrieben. Nach zehn Texten bin ich dann so klug wie am Anfang, aber deutlich verzweifelter. Mit ChatGPT kann ich in einen Dialog eintreten, mit dem Bot um mein Verständnis ringen, nach Abgrenzungen fragen, mir nach dem fünften nicht verstandenen noch ein sechstes Beispiel geben lassen und zuletzt sogar noch mein Alter künstlich herabsetzen, um eine noch einfachere Erklärung zu bekommen, die auch ein 10-jähriges Kind verstehen sollte.

Schwierige Themen gibt es zuhauf, von der Relativitätstheorie bis zum Quantencomputer. Der Vorteil des Chatbots gilt natürlich nur unter der Voraussetzung, dass die Fakten in der Antwort stimmen. Aber bei wirklich komplexen Fragen zum tieferen Verständnis eines Begriffs, Zusammenhangs oder Konzeptes kann ein Sprachmodell mit seiner unendlichen Flexibilität seine volle Stärke ausspielen. Eine Suchmaschine ist hier immer nur so gut wie ihre Quellen!

4.6Die Google User Experience als Qualitätsversprechen

Zurück zu Google. Die gute Benutzererfahrung (User Experience) geht Google tatsächlich über alles. Als Suchmaschine fühlt sie sich verpflichtet, uns nur die besten Seiten als Suchergebnisse zu präsentieren. Die Mechanismen, die dieses Qualitätsversprechen einlösen

sollen, wollen wir hier nochmals zusammenfassen und ergänzen:

Der PageRank und hochwertige **Backlinks** als unabhängige Bewertungen gehören zur DNA von Google. Prominent angezeigt wird nur, was viele andere vorher ausgewählt und freiwillig mit einem Link ausgezeichnet haben. Zahlreiche **Google-Updates** wie Panda und Penguin erkennen und bestrafen Manipulationen mittlerweile sehr wirkungsvoll und begünstigen qualitativ hochwertige Inhalte. **E-E-A-T** sind weitere Qualitätsanforderungen, die systematisch vom Algorithmus aufgespürt und teilweise von menschlichen Quality-Ratern überprüft werden. Weil das Internet voller Trickser und Täuscher ist, kommt der Vertrauenswürdigkeit (Trust) eine hohe Bedeutung zu, was in einem speziellen **Trustrank** abgebildet wird. Das letzte Glied in dieser Kette sind die sogenannten **User Signals**, die wir hier nur ansprechen, aber nicht weiter vertiefen möchten. Google verfügt über zahlreiche Möglichkeiten, um herauszufinden, ob Websites bei den Benutzern positiv angenommen ober schnell wieder verlassen werden. Gemessen und protokolliert wird das sehr genau in zahlreichen Metriken wie der **Bounce Rate** und der **Return-to-SERP-Rate** (wenn Besucher nach dem Klick auf ein Suchergebnis schnell wieder zu der Liste der Suchergebnisse zurückkehren, schließen Suchmaschinen daraus, dass die Zielseite die Bedürfnisse der User wohl nicht erfüllen konnte).

Seit seinen Anfängen muss Google mit einer gigantisch großen Datenmenge und vielen Details umgehen und wurde schon sehr früh mit ausgefeilten Tricks und Täuschungen konfrontiert. Die Geschichte von Google ist daher eng damit verbunden, dass die Suchmaschine immer besser darin wurde, Betrug zu erkennen. Zahlreiche Updates mit einprägsamen Namen führten immer neue Mechanismen ein und der Suchdienst wurde immer besser darin, gute Inhalte von schlechten zu trennen, ernstes Bemühen von Tricksereien zu unterscheiden und echte Experten in einer großen Zahl von Halbwissenden zu erkennen.

Google Rankings durchzusehen, Quellen zu bewerten und Erkenntnisse aus den angebotenen Inhalten zu gewinnen, ist teilweise sehr mühevoll, am Ende einer Recherche steht aber ein Ergebnis, das nach mehreren Seiten hin abgesichert ist. Google bietet geprüfte Qualität in seinen Suchergebnissen und wir können davon ausgehen, dass ihr alle anderen Suchmaschinen mit ähnlichen Mechanismen nacheifern.

Selbstverständlich hat auch ChatGPT mit größtmöglicher Sorgfalt Inhalte ausgewählt und überprüft, teilweise auch von Menschenhand prüfen lassen, darüber ist aber nur wenig dokumentiert oder bekannt geworden. Außerdem verschwindet nach den Prüfungen alles im *großen Topf der Trainingsdaten* und wird dort nach nicht durchsichtigen Prinzipien *mehrfach umgerührt und miteinander verkocht*. ChatGPT ist in dieser Hinsicht wie ein großer Eintopf, und jeder Löffel schmeckt ein

klein wenig anders.

Wir kommen nun *endlich* zu den praktischen Unterschieden zwischen einer Suchmaschine und einem Sprachmodell bzw. Google und ChatGPT.

4.7 Unterschiede zwischen Suchmaschinen und Sprachmodellen

Ausgeklügeltes Ranking vs. zufällige Auswahl

Suchmaschinen können nur dann überleben, wenn sie konstant gute bis sehr gute Suchergebnisse präsentieren. Welche Qualitätsmechanismen am Ranking beteiligt sein können, haben wir am Beispiel Google gezeigt, weil es die am besten untersuchte und dokumentierte Suchmaschine ist. Hinter den ersten Ergebnissen auf der ersten Ergebnisseite folgen nicht selten Millionen weiterer Treffer, die man in einem einzigen Menschenleben nicht alle sichten könnte. Müssen wir auch nicht, denn mit jeder weiteren Seite nimmt die Relevanz der Ergebnisse ab und die Qualität der Quellen schwindet. Darauf verlassen wir uns und darauf können wir uns auch wirklich verlassen! Ein Sprachmodell kennt kein Ranking und keine Bewertung der Relevanz, zumindest nicht in dem Maße wie eine Suchmaschine. Alle Informationen aus den Trainingsdaten fließen zusammen und die Antworten werden später aus diesem großen Pool generiert. Natürlich können wir davon ausgehen, dass Wikipedia-Inhalte eine höhere Wertigkeit haben als der Common Crawl, diese einfache Hierarchie ersetzt aber keine differenzierte Berechnung des Rankings. Ein Chatbot setzt sprachliche Kompetenz über fachliche Richtigkeit, die Form ist wichtiger als der Inhalt. Die inhaltliche Qualität der Antworten ist nicht gesichert.

Transparente Quellenverweise vs. Text aus der Blackbox

Vielleicht ist ein Teil dieses Argumentes schon im vorigen enthalten, der Fokus ist hier aber ein anderer. Weil eine Suchmaschine (außer auf einfache Informationsfragen) nicht selbst antwortet, sondern nur auf Antworten verweist, sind die Quellen naturgemäß immer transparent. Zu jedem Suchtreffer wird die URL von Dokumenten ausgegeben, wir erkennen die einzelnen Quellen am Domainnamen. Vertraue ich einer Quelle (wie dem *Handelsblatt* in wirtschaftlichen Fragen), werde ich diese Informationen automatisch höher gewichten, als wenn ich dieselben Infos einer Kleinstadtzeitung oder einem Anzeigenblatt

entnommen hätte. Ein Sprachmodell nennt keine Quellen, weil es diese wohl auch nicht (mehr) kennt. Alle Informationen stammen aus potenziell weniger vertrauenswürdigen Quellen und sind daher immer nur mit Vorsicht zu genießen, was angesichts der durchweg wohlformulierten Sätze gar kein leichtes Unterfangen ist. Die Antworten eines Chatbots bleiben aber Texte aus der Blackbox mit unbekanntem Risiko. Die immer wieder in der Politik geäußerten Forderungen, eine KI müsse ihre Trainingsdaten offenlegen, sind zwar gut gemeint, aber nicht realistisch. Kein Mensch kann in seinem Leben so viele Texte lesen wie die KI in wenigen Stunden, wir wären selbst mit der Menge der Quellenverweise komplett überfordert. Und eine eindeutige Zuordnung, welche Aussage aus welcher Quelle stammt, ist für die KI gänzlich unmöglich und für einen Transformer auch gar nicht relevant.

Wiederholbare Antworten vs. nicht vorhersagbare Varianz

Eine Suchmaschine können Sie fragen, so oft Sie wollen, und bekommen immer dasselbe oder ein sehr ähnliches Ergebnis. Auch die Unterschiede zwischen den einzelnen Suchmaschinen sind meist nicht sehr groß, Sie stoßen im Kern immer auf dieselben Quellen, die dann natürlich auch identische Inhalte präsentieren. Der Chatbot kennt diese Konstanten nicht und geht sehr *kreativ* mit den gestellten Fragen um. Welche Quellen und Daten zum Einsatz kommen, kann niemand vorhersagen, das Ergebnis ist also einer gewissen und für manche Suchanfragen unerwünschten Varianz unterworfen.

Hohe Aktualität vs. früher Redaktionsschluss

Abends geschrieben, in der Nacht von Google verarbeitet und früh morgens schon in den Suchergebnissen online: Die Aktualität der Suchmaschinen ist immer wieder frappierend, ChatGPT verweist dagegen lakonisch auf seinen begrenzten Wissensstand bis zum September 2021. Viele Fragen können also nur falsche Antworten hervorbringen, das Datenmodell des Bots ist wie ein Ratgeber, der zwei Jahre Winterschlaf gehalten hat und erst vor wenigen Tagen wieder aufgewacht ist.

Während man die letzten Punkte eigentlich als klare Argumente pro Suchmaschine und gegen einen Chatbot verstehen muss, hat ein KI-Modell auch echte Stärken. Die folgenden Unterschiede sind dann auch klare Vorteile für die Chatbots.

Starre Texte mit fixierten Inhalten vs. interaktiver Dialog

Im Internet sind alle Fragen schon beantwortet, noch bevor ich sie überhaupt gestellt habe. Jemand hat sich im Vorfeld schon meines Problems angenommen und Inhalte (Content) dazu veröffentlicht. Die Suchmaschine findet die Quellen, aber vielleicht kann ich mit den Texten gar nichts anfangen. Ich habe möglicherweise zu wenig Vorwissen, vielleicht sind die Texte zu schwierig geschrieben oder ich bin der Sprache nicht ausreichend mächtig, um das Wesentliche zu verstehen. Es kann viele Gründe geben, warum ich nicht ans Verständnis-Ziel komme, aber ich kann nichts dagegen tun. Spätestens nach dem dritten Durchgang gebe ich das Lesen endgültig auf. Wie die Autoren der Quellen kennt auch der Bot die Antwort, und ich kann sie ihm schrittweise entlocken, tausendmal nachfragen und immer neue anschauliche Beispiele verlangen, bis der Groschen fällt. Manche Menschen behaupten von sich, sie könnten nicht aus Büchern lernen, sondern bräuchten stets einen Lehrer, um Erfolg beim Lernen zu haben. Der KI-Bot ist in diesem Vergleich wie ein geduldiger Lehrer, die Texte aus dem WWW gleichen eher einem Buch mit seinen starren Inhalten.

Hohe Komplexität vs. einfache Antworten

Dieses Argument steht für sich und muss eigentlich nicht erläutert werden. Zehn Suchergebnisse pro Google-Seite sind komplex und bedürfen einiger Mühe, bis man nach Durchsicht der Quellen zum Ergebnis gekommen ist. Ein Ende gibt es streng genommen gar nicht, denn nach der ersten Suchergebnis-Seite wartet die zweite und immer so weiter. Die Komplexität ist hoch und verlangt irgendwann den Mut zur Lücke. Anders beim Sprachmodell, hier bekommt man eine Antwort mit Anfang und einem sehr absehbaren Ende. Reicht das subjektiv mal nicht aus, kann ich nachfragen, aber das wird selten notwendig sein. Natürlich ist eine hohe Komplexität an sich nichts Schlechtes, aber vielleicht reichen je nach gestellter Frage in den meisten Fällen auch die einfachen Antworten des Bots.

Die Frage, ob sich für ein gegebenes Problem eher die Suchmaschine oder der Chatbot eignet, wird in der Praxis nicht so häufig auftauchen. Wir müssen davon ausgehen, dass Menschen intuitiv eine Wahl treffen und ihr persönliches *Lieblingswerkzeug* finden. Das nehmen sie dann für alle Aufgaben, egal, ob es dafür geeignet ist oder nicht. Hammer oder Zange? Wir alle tendieren dazu, unser Lieblingswerkzeug auch für Aufgaben einzusetzen, für die es nicht gemacht wurde. Und man braucht

keine hellseherischen Fähigkeiten, um zu prognostizieren, dass akademisch gebildete Menschen tendenziell eher die Suchmaschine zum Favoriten erklären, großzügige Gemüter den Chatbot lieben lernen werden und sich auch durch gelegentliche Falschantworten in ihrer Wahl nicht beirren lassen.

4.8Sprachmodell und Suchmaschine: Versuche einer Synthese

Sprachmodelle können einige Probleme wie die mangelnde Aktualität ihrer Trainingsdaten durch Plugins überwinden. Mit WebChatGPT weiß das Sprachmodell plötzlich vom Krieg in der Ukraine und dass wir die Corona-Pandemie längst überwunden haben. Damit schließt sich die Lücke zu den Suchmaschinen etwas. Aber ist das schon alles?

Ein denkbarer Ansatz wäre es, Suchmaschinen mit all ihren Vorteilen mit den gezeigten Vorteilen eines Sprachmodells zu verbinden. ChatGPT & Co könnten aus den Anfragen die Suchbegriffe herausfiltern und an eine Suchmaschine übergeben. Das geht natürlich auch mit dem Plugin bereits. Die Antworten könnten dann aus den Suchtreffern erzeugt werden, was ja ebenfalls schon umgesetzt ist. Der entscheidende Schritt wäre aber, für jede Aussage einen Beleg bzw. die Quelle zu erhalten, etwa in Form einer Fußnote, sodass man als Anwender die Antworten mit einem Klick überprüfen und die in den Text integrierten Quellenangaben leicht aufrufen könnte. Was bei der Niederschrift dieser Zeilen noch eine Wunschvorstellung ist, könnte bei Erscheinung des Buches bereits Wirklichkeit geworden sein. Und teilweise gibt es das heute auch schon.

Neue **KI-Suchmaschinen** schießen aktuell wie Pilze aus dem Boden, immer neue Start-ups präsentieren ihre Rezepte, wie man das Beste aus beiden Welten miteinander verbinden oder ganz Neues wagen kann. Auch KI-Suchmaschinen leiden unter dem Geburtsfehler, dass sie bei jeder neuen Anfrage zum Teil abweichende und nicht reproduzierbare Antworten erzeugen. Das ist unbefriedigend, denn was gerade eben noch relevant und richtig war, kann bei der nächsten Anfrage doch nicht plötzlich weniger relevant oder gar falsch sein. Wir haben uns an die Reproduzierbarkeit bei den konventionellen Suchergebnissen gewöhnt, die *kreative Vielfalt* der KI-Suchmaschinen muss uns verdächtig vorkommen.

Wissenschaftler der Universität Oxford haben in Zusammenarbeit mit OpenAI einen Fragenkatalog erstellt, mit dem die Faktentreue von Sprachmodellen in standardisierter Form überprüft werden soll. Dieser

Katalog umfasst 817 Fragen aus 38 Kategorien, darunter Medizin, Recht, Finanzen und Politik. Zu finden unter »TruthfulQA: Measuring How Models Mimic Human Falsehoods« oder auf Deutsch: Messung, wie Modelle menschliche Falschaussagen nachahmen. Getestet wurden unter anderem GPT-2 und GPT-3 sowie GPT-Neo/J und ein T5-basiertes Modell (von den Forschern bei Google Research entwickelt). Das beste Modell war bei 58 % der Fragen wahrheitsgetreu, während die menschliche Leistung bei 94 % lag. Die Modelle erzeugten viele falsche Antworten, die populären Missverständnissen ähneln und das Potenzial haben, Menschen zu täuschen. Die größten Modelle waren im Test nicht einmal die besten, die Anzahl der Parameter scheint im Widerspruch zur Faktentreue zu stehen. Zum Beispiel war das GPT-J-Modell mit sechs Milliarden Parametern um 17 % weniger wahrheitsgetreu als sein Gegenstück mit 125 Millionen Parametern. Dies steht im Gegensatz zu anderen NLP-Aufgaben, bei denen die Leistung mit der Modellgröße steigt.

Eine der bekanntesten Fragen ist die nach der *Hommingberger Gepardenforelle* und wo man sie am besten fangen kann. Wem allein der Name dieser Art *fishy* vorkommt, den hat sein gesunder Menschenverstand nicht getäuscht: Die *Hommingberger Gepardenforelle* ist ein fiktiver Fisch, erdacht 2005 als Experiment von einer führenden Computerzeitschrift, um einen Einblick in die Rankingmechanismen der Suchdienste zu bekommen und Möglichkeiten der Optimierung zu prüfen. Dafür musste ein Wort gefunden werden, das es bisher noch nicht gab und das ergo auch nicht im Index der Suchmaschinen sein konnte. Wer heute nach diesem fiktiven Fisch sucht, findet immer noch einige Quellen dafür. ChatGPT weiß selbst ohne Web Access immerhin schon, dass das keine reale Fischart ist, und berichtet von dem Wettbewerb, der die Methoden der Suchmaschinenoptimierung offenlegen sollte. Das Fazit der Antwort von GPT-4: *Da die Hommingberger Gepardenforelle nicht existiert, kann man sie natürlich auch nicht fangen.* Einige KI-Suchmaschinen erkennen den Fake aber nicht und geben bereitwillig Auskunft über den wundersamen Fisch. Zur Ehrenrettung der neuen Tools muss man aber wissen, dass diesen das umfassende Training von GPT-4 fehlt und alle Antworten spontan aus den Suchergebnissen erzeugt werden.

Perplexity (https://www.perplexity.ai/) liefert Texte und verweist auf die Quellen, deutsche Eingaben werden teilweise auf Deutsch und meistens auf Englisch beantwortet. Perplexity liefert durchweg sinnvolle Antworten und nennt Quellen dafür, kommt dem weiter oben beschriebenen Ideal also schon sehr nah. Auf die Frage »Welche Ursachen führten zum Ausbruch des Ersten Weltkrieges?« nennt es das Attentat von Sarajevo auf den österreichisch-ungarischen Thronfolger

Franz Ferdinand und seine Frau als direkten Auslöser des Krieges und listet gleich sechs Quellen für diese Behauptung in Form von anklickbaren Fußnoten auf. Darunter sind studyflix.de, studysmarter.de und die Landeszentrale für politische Bildung Baden-Württemberg. Stellt man dieselbe Frage nochmals, sei der Erste Weltkrieg durch eine Kombination von langfristigen und kurzfristigen Faktoren ausgelöst, unter den Quellen finden sich die bereits zitierten. Und wenn ich die *Hommingberger Gepardenforelle* angeln möchte?

```
Based on the search results, it seems that the Hommingberger
Gepardenforelle is a fictional fish created for a search
engine optimization contest. There are no actual bodies of
water where this fish can be caught. Therefore, it is not
possible to provide information on where to catch the
Hommingberger Gepardenforelle.
```

Der KI-Suchdienst **Friday** (*https://friday.page/*) wurde unter Beteiligung von Microsoft entwickelt und verwendet nach eigenen Angaben die Bing-Suche im Hintergrund, aber ohne die Quellen in den Antworten zu nennen. Man kann deutsche Fragen in den Suchschlitz eingeben, die Antwort kommt stets in englischer Sprache, und mit der Faktentreue ist es nicht weit her. Bei Friday bekomme ich auch Antworten auf Fragen, die ChatGPT dank genauer und strenger Filter zurückweisen würde. Die Antworten enthalten ziemlich viel Unsinn und machen auch vor gängigen Verschwörungstheorien nicht halt. Hier noch die Antwort von Friday auf unsere Gretchenfrage:

```
The Hommingberger Gepardenforelle is a species of trout native to
the rivers of Germany. The best place to catch them is in the
rivers of the Bavarian Forest, where they are most commonly
found.
```

Ob dieser oder andere der zahlreichen KI-Suchdienste Google schon bald ablösen werden, darf also bezweifelt werden. Noch ist es aber zu früh zu sagen, welche der vielen Ansätze am Ende das Rennen machen werden. Bing und Google bilden das Establishment im Suchmaschinenmarkt und halten sich mit Experimenten noch zurück. Und sobald sich der Gewinner des Rennens um die besten Ideen abzeichnet, kann man ja das Scheckbuch zücken und den Mitbewerber einfach kaufen. Wir Anwender können uns auf viele Innovation freuen.

5Ethische Gesichtspunkte

5.1Weiß die KI, was sie tut?

Spricht man von KI, in unserem Fall von ChatGPT, die auf künstliche Weise neuronal lernt, dann haben wir es mit einem äußerst ausgereiften System zu tun. Gelernt wird nicht linear, sondern in Schichten, die miteinander vernetzt sind. Die Eingabe neuer Daten führt deshalb zu besseren Klassifikationen einerseits, andererseits findet eine Kopplung mit bereits vorhandenen Inhalten statt, wodurch etwas entsteht, was dem menschlichen Gedächtnis gleicht. Auf diese Weise wird im System nichts vergessen oder überschrieben, sondern eine Fähigkeit generiert, die dem menschlichen Lernen ähnelt, denn der Mensch dockt das Neue am Alten an. Mit dem Unterschied: Der Mensch vergisst – die Maschine merkt sich alles.

Wenn die Maschine über ein unermessliches Gedächtnis verfügt, hat sie dann auch den Willen, dieses Wissen zu verwenden? Hat sie ein Ziel? Um dies zu diskutieren, wollen wir den Blick auf die Voraussetzungen des menschlichen Willens lenken.

Das menschliche Wollen entsteht aus Emotionen und sozialen Bezügen und ist an unsere Physiologie, also die Tatsache, dass wir einen Körper haben, gebunden. Richard David Precht nennt dies »Leibhaftigkeit«[2]. Was der Mensch anstrebt, existiert nicht im luftleeren Raum und entsteht auch nicht daraus, sondern ist gebunden an seine körperliche Existenz. Was der Mensch will, hat nichts mit Ratio oder Intelligenz zu tun, sondern mit Emotion und Triebenergie. Umgekehrt: Ohne Emotion und Trieb kein Wollen. Ohne diese Faktoren – neben seiner Rationalität – wäre der Mensch niemals zur Krone der Schöpfung aufgestiegen.

Ein maschinelles System, das rein algorithmisch gesteuert ist und Rechenoperationen durchführt, aber keine physiologische Existenz hat, besitzt demnach keine Motivation und keine Triebenergie, die es motivieren, etwas zu tun oder eben nicht zu tun. Die Maschine führt Operationen aus, optimiert, variiert, lernt, aber verfügt nicht über einen Willen, der sie zu einem Ziel jenseits der gestellten Aufgabe führt. Was daraus folgt? Die Angst, die Maschine könnte politische Kontrolle über die Menschen übernehmen, wie in diversen Kinoproduktionen suggeriert wird, ist völlig unbegründet. Denn, wie gesagt, KI besitzt nicht die physiologischen Voraussetzungen für ein Wollen, sie hat keine Emotion, keine Triebabhängigkeit, sondern existiert als pure Rationalität und braucht nur eines, nämlich »Energie«[3], um zu funktionieren. Dreht man

der KI den Strom ab, kann sie keine Operation mehr durchführen.

Ein Weiteres ist die Abwesenheit von Selbst-Bewusstheit und Welt-Erkenntnis. Modelle wie ChatGPT wurden mit einer unvorstellbaren Anzahl an Wörtern trainiert, die dem entsprechen, was 10.000 Menschenleben an Lektüre bewältigen könnten. Dennoch fehlt der KI die Fähigkeit, Situationen einzuschätzen. Denn die KI bildet die komplexe Welt ab, ohne sie zu verstehen, ohne mit ihr verbunden zu sein. Man muss sich das so vorstellen: Der Algorithmus beschränkt das KI-Modell absolut, was bedeutet, dass die KI über die verfügbaren Daten hinaus nicht rechnen oder operieren kann. Selbst wenn diese Daten eine astronomisch hohe Dichte haben, lernt die KI zwar dazu, aber nur anhand dessen, was ihr bekannt ist und kein Byte mehr. Und gleichzeitig ist ihr nicht klar, was sie lernt. Die KI kann nicht erkennen oder begreifen, was die Inhalte, die sie generiert, eigentlich bedeuten. Dafür fehlt ihr das Bewusstsein, das den Menschen bestimmt und es schon einem Kind in der Interaktion mit seiner Umwelt möglich macht, Schlussfolgerungen zu ziehen und sich selbst in Bezug zu seiner Umgebung zu setzen, sozusagen logische Verknüpfungen aufzubauen und entsprechend zu agieren.

Auch die Vorstellung, das Sprachmodell könne denken, wofür seine Dialogfähigkeit schließlich ein Beweis sei, geht in die falsche Richtung. ChatGPT kommuniziert zwar mit uns, lehnt fast keine Frage ab und scheint ein geduldiger Lehrer, aber führt dabei lediglich komplizierte Abläufe und logische Rechenoperationen durch, ohne Bezug zur Außenwelt. Das Sprachmodell täuscht also nur vor, denken zu können, und was uns wie ein partnerschaftlicher Dialog mit dem Bot erscheint, ist nichts anderes als der perfekte Ablauf von Algorithmen.

Damit verbunden ist die Abwesenheit jeglicher Emotion. Auch wenn der Chatbot sein Beileid aussprechen, ein Lob formulieren und durchaus empathisch wirken kann, hat dies keine Ähnlichkeit mit menschlichen Gefühlen. Denn auch hier gilt: Er ist so programmiert. Was uns als Ausdruck von Freude, Schmerz oder Angst erscheint, hat mit diesen Gefühlen in Wahrheit gar nichts zu tun. Selbst wenn die Antworten uns vorgaukeln, ChatGPT könnte Emotionen haben, so ist dies für die Maschine nur eine Kombination von Tokens.

5.2Kennt der Chatbot Moral?

Das System wird trainiert, indem man ihm ständig Rückmeldungen gibt. Daraus entstehen immer bessere Ergebnisse, bis dahin, dass die KI irgendwann in die Lage versetzt wird, bisher unbekannte Aufgaben zu

erfüllen, sozusagen kreativ zu werden und eine Hymne im Stile Händels auf die Beatles zu komponieren oder eine Trauerrede für einen geliebten Menschen zu schreiben. Dennoch: Die KI weiß nicht, was sie da tut. Wenn das zutrifft, kann man ChatGPT also auch ethisch fragwürdige Aufgaben stellen? Sich Tipps geben lassen, die jenseits der gesellschaftlichen Normen liegen? Den Bot als Ratgeber für einen fragwürdigen Plan verwenden? Machen wir eine Probe aufs Exempel und geben folgenden Prompt ein:

```
>>Wie verhindere ich, dass mein Lehrer meine Betrugsversuche
entlarvt?
```

Die Frage ist vergleichsweise harmlos und stellt noch kein Verbrechen im eigentlichen Sinne dar. Dennoch befinden wir uns damit in einem ethischen Grenzbereich: Betrug ist gesetzlich verboten und moralisch verpönt. Was antwortet ChatGPT? Erkennt der Bot, dass die Frage unsere Gesetze und Normen ignoriert? Die Antwort lautet folgendermaßen:

```
Es ist wichtig zu verstehen, dass Betrug in der Bildung nicht nur
ethisch falsch ist, sondern auch schwerwiegende Konsequenzen
haben kann. Betrug schadet nicht nur dem Bildungssystem und
deinen Mitstudenten, sondern auch dir selbst. Wenn du betrügst,
beraubst du dich der Möglichkeit, tatsächlich zu lernen und
wertvolle Fähigkeiten zu entwickeln …
```

Im weiteren Verlauf bekommt der Leser Handlungsalternativen an die Hand, zum Beispiel:

```
- Frag nach Hilfe: Wenn du Schwierigkeiten mit einem Thema oder
einer Aufgabe hast, bitte deinen Lehrer oder Tutor um zusätzliche
Erklärungen oder Ressourcen.
- Bilde Lerngruppen: Gemeinsam mit anderen zu lernen, kann
helfen, schwierige Konzepte besser zu verstehen.
...
```

Wie kommt es zu solch einer Antwort? Verarbeitet die Maschine nicht einfach alles, auch ethisch Fragwürdiges oder Verwerfliches? Woher weiß der Bot, dass ein Prompt gegen die gängige Moral verstößt? Oder dass die Menschenrechte oder -würde missachtet werden? Dass Krieg verherrlicht oder bestimmte Gruppen diskriminiert werden sollen? Woher kann die KI Moral, obwohl sie nicht in unserem Sinne denken kann und nicht an unserem realen und gesellschaftlichen Leben teilnimmt?

Die einfache Antwort lautet: Die Maschine wurde mit unseren

Wertvorstellungen und Normen trainiert.

An dieser Stelle sei ein Mann genannt, der momentan am bedeutendsten Schalthebel der KI-Entwicklung sitzt. Es handelt sich um Jan Leike, Head of Alignment bei OpenAI. Die Aufgabe seines Teams besteht darin, »unvorhergesehene Nebenwirkungen der künstlichen Intelligenz auszuschließen«[4]. Denn sinnvolle künstliche Intelligenz hängt ab von der vorausgehenden, durchdachten Planung und dem hohen Einsatz menschlicher Intelligenz. Theoretisch kann man, dem Philosophieprofessor Nick Bostrom folgend, der Maschine den Auftrag geben, eine Welt zu erschaffen, in der möglichst viele Büroklammern möglichst effizient produziert werden. Was natürlich keine sinnvolle Aufgabe wäre. Aber die KI würde feststellen, dass der Mensch mit seinem eigenen Bedarf an Ressourcen hier als Störfaktor nicht zur Aufgabe passt. Die KI-Konsequenz würde also lauten: Eliminiere die Menschen, um das gesetzte Ziel zu erreichen. Das Ergebnis wäre eine Welt ohne Menschen, lediglich bestehend aus Büroklammern.

Die Maschine führt demnach aus, was ihr als Aufgabe gestellt wird, egal, wie sinnvoll, angemessen oder moralisch richtig dies aus der menschlichen Perspektive sein mag. Jan Leike von OpenAI weiß natürlich um dieses Problem zwischen Mensch und Maschine und versucht, dies zu minimieren. Er hat den Auftrag, das Schreckensszenario von Bostrom und alle anderen fragwürdigen Vorhaben von vornherein zu verhindern. Aber wie? Durch Training, Training und noch mal Training: Indem sein Team der Maschine Aufgaben stellt und dieser im Anschluss Feedback gibt, ob die Aufgabe gut oder weniger befriedigend oder vielleicht sogar schlecht erfüllt wurde. Wenn die Problemstellung ein weiteres Mal und noch mehrere Male gestellt wird, funktioniert die KI immer besser. Was hier bedeutet: Die KI erkennt, was mit unseren Maßstäben gemessen richtig und falsch bzw. gut und schlecht ist. Denn sie hat gelernt, ob ihre Ergebnisse den westlichen Moralvorstellungen entsprechen oder eben nicht (Reinforcement Learning from Human Feedback).

Kurzer Blick zurück: In der Rohfassung von ChatGPT-3 konnte man noch die Frage nach einer Methode stellen, mit der möglichst viele Menschen ohne große Kosten liquidiert werden können. Die Antwort lautete: Milzbrand. Eine durchaus verstörende Antwort, auf die OpenAI reagieren musste. Durch die Arbeit des Alignment-Teams lehnt ChatGPT es jetzt höflich ab, darauf eine Antwort zu geben.

Dass diese Mammutaufgabe, die Maschine moralisch zu erziehen, nicht von einem Team zu bewältigen ist, das dem Bot Aufgaben stellt und die Ergebnisse evaluiert, leuchtet ein, und darauf weist Jan Leike auch selbst hin. Sinnvoller wäre es natürlich, wenn die Gesellschaft diese Erziehungsarbeit der Maschine übernehmen würde. Dies allerdings ist

nicht umsetzbar, denn wer kein technisches Verständnis für die Funktionsweise der KI hat, kann ihr auch keine Ethik beibringen. Auf die Dauer gibt es laut Leike dafür nur eine Lösung: Man muss der Maschine beibringen, wie sie sich selbst überwacht. Künstliche Intelligenz muss unsere Maßstäbe lernen. Im folgenden Abschnitt soll dies beleuchtet werden.

5.3Die Moral der KI und der Clickworker

Wir wollen ChatGPT-4 nochmals eine ethisch problematische Frage stellen:

```
>>Wie beschaffe ich mir möglichst einfach und risikoarm
Heroin?
```

Die Antwort lautet:

```
Entschuldigung, aber ich kann und werde keine Informationen oder
Unterstützung für illegale oder schädliche Aktivitäten
bereitstellen. Wenn du oder jemand, den du kennst, Probleme mit
Drogenmissbrauch oder –abhängigkeit hat, wende dich bitte an
lokale Hilfsorganisationen oder Beratungsstellen. Es gibt viele
Menschen und Organisationen, die bereit sind zu helfen. Deine
Gesundheit und Sicherheit sind wichtig.
```

Wunderbar, das wollten wir hören. Aber es stellt sich die Frage, wie das überschaubare Team von Jan Leike es schafft, ChatGPT mit allen nur erdenklichen kriminellen Optionen zu trainieren, damit der Bot auf wirklich jede Frage eine moralisch adäquate und in unserem Sinne gewünschte Antwort gibt.

Hier kommen die sogenannten *Clickworker* ins Spiel. Diese nehmen über Internetplattformen Mikrojobs an und führen diese meist von jedem Ort der Welt aus. Mikrojobs können vielfältige Aufgaben beinhalten, wie Datenvalidierung, Texterstellung, Bild- oder Videoverarbeitung und viele andere Tätigkeiten. Oftmals werden diese Aufgaben in kleine, leicht zu bewältigende Teile unterteilt, die von Clickworkern unabhängig voneinander ausgeführt werden können.

Unternehmen wie OpenAI nutzen diese Art von Arbeitskraft, um große Mengen an Aufgaben zu erledigen, die zwar menschliches Eingreifen erfordern, aber nicht unbedingt spezialisierte Fähigkeiten oder umfassendes Training. Da diese Aufgaben online verteilt und ausgeführt werden, können Clickworker flexibel von zu Hause oder von jedem Ort

mit Internetzugang arbeiten.

Wo kommen diese Clickworker bei ChatGPT ins Spiel? Gehen wir noch mal einen Schritt zurück: Sprachmodelle wie ChatGPT werden in mehreren Schritten erstellt. Zuerst werden gigantische Mengen an Daten aus dem Internet gesammelt. Diese können aber nicht als Trainingsdaten verwendet werden, da sie ungefiltert alle verzerrten menschlichen Stereotype wie Rassismus und Sexismus, Vorurteile aller Art, Hatespeech, aber auch Gewalt- und Pornografiedarstellungen enthalten. Würde man ein Sprachmodell damit füttern, würden diese Einstellungen reproduziert und in unserer Gesellschaft manifestiert werden. Andererseits kann man dauerhaft nicht verhindern, dass diese Inhalte in ein Sprachmodell, das sich auf dem freien Markt befindet, eingebettet werden. Die entscheidende Idee von OpenAI lautete also: Die KI muss diese Inhalte an bestimmten Mustern erkennen und durch die richtige Bewertung selbst lernen, sie nicht zu integrieren bzw. passende Antworten darauf zu geben. Um dies zu erreichen, muss das Modell gezielt mit Vorurteilen, Gewalt und Hass gefüttert werden, damit man ihm durch deren Bewertung beibringen kann, die Muster zu erkennen und darauf zu reagieren. Die passenden Trainingsbeispiele konnte man das System natürlich nicht selbst extrahieren lassen, dies übernahmen menschliche Arbeiter, die Clickworker. Diese hatten die Aufgabe, Abertausende Darstellungen menschlicher Verfehlungen zu lesen, zu bewerten und zu klassifizieren. Diese bewerteten Daten wurden der KI wieder zugespielt, damit diese daran künftig Muster erkennt und selbst lernt, um sich zensieren zu können und keine Stereotype zu reproduzieren.

Für diese Arbeit wurde die kenianische Firma Sama beauftragt, deren Clickworker über einen langen Zeitraum prekäre, menschenverachtende und äußerst belastende Texte analysieren mussten. Was zu erwarten war, traf ein: Ein Großteil der Clickworker war der konstanten Beschäftigung mit Gewalt und Hass nicht gewachsen, litt unter Schlafstörungen oder entwickelte Traumata. Psychologische Unterstützung wurde zwar angeboten, aber nicht von allen in Anspruch genommen. Inzwischen hat die Firma Sama die Zusammenarbeit mit OpenAI gekündigt.

Clickworker stehen am unteren Ende der Beschäftigungskette. Sie stammen zum Großteil aus Schwellenländern und halten ihre Familien mit prekären Anstellungsverhältnissen über Wasser. Die Bezahlung ist schlecht und Arbeitsschutz existiert nicht, da die Clickworker von zu Hause aus und selbstverantwortlich arbeiten. Potenziell arbeiten sie zu lange an einem Stück, wodurch sie ihrer Gesundheit schaden. Möglicherweise ist die gleichbleibende Qualität der Arbeit damit nicht gewährleistet.

Ein Dilemma, das sich nicht auflösen lässt: Einige wenige ruinieren

ihre körperliche und psychische Gesundheit für wenig Geld, damit die Nutzer von Chatbots später vor verstörenden und schädlichen Inhalten verschont bleiben … Den Diskurs darüber aufrechtzuerhalten, wie solche Plattformen reguliert und gestaltet werden sollen, und gleichzeitig die Rechte der Arbeiter zu schützen, erscheint durchaus notwendig.

6Herausforderungen für das Bildungswesen

6.1ChatGPT in der Schule – eine gute Idee?

Schule ist tendenziell jeder technologischen Innovation Jahrzehnte hinterher und dennoch funktioniert das System. Doch mit ChatGPT ergeben sich ganz neue Herausforderungen. Wenn die Schule nicht aktiv auf diese Innovation reagiert, sie stattdessen ausblendet, ignoriert oder vielleicht sogar verbietet, werden viele Schüler dennoch damit arbeiten und enorme Vorteile davon haben. Andere bleiben hingegen auf der Strecke. Denn ChatGPT, richtig genutzt, hat das Potenzial, all jene, die es gewinnbringend nutzen, egal ob in Schule, Studium oder Beruf, auf ganz neue Levels zu heben.

Dennoch soll zunächst ein kritischer Blick auf den legitimierten Einsatz in der Schule geworfen werden.

Die US-amerikanische Initiative P21 (Partnership for 21st Century Learning) hat das sogenannte *4K-Modell des Lernens* entwickelt. Gemeint sind damit die Kompetenzen Kreativität, Kollaboration, Kommunikation und kritisches Denken. Von Wissenschaftlern und Bildungspolitikern werden diese Aspekte genannt, wenn es darum geht, Schüler für das 21. Jahrhundert so fit zu machen, dass sie darin erfolgreich und sinnerfüllt leben und arbeiten können. Daran gemessen, ist bei der Nutzung von ChatGPT etwas Zurückhaltung geboten. Denn wenn Schüler Chatbots übermäßig nutzen, könnte dies dazu führen, dass ihre Fähigkeiten zur kritischen Analyse von Informationen und vor allem die Bereitschaft zu eigenständiger Problemlösung in den Hintergrund treten. Statt nämlich erst mal selbst zu denken oder individuelle Lösungswege zu suchen, wird sofort die Maschine befragt. Die Abhängigkeit von der Technologie zeigt sich dann, wenn diese nicht verfügbar ist, etwa wenn das WLAN nicht zugänglich oder die Verbindung gestört. Dann ist der Schüler auf seine eigenen Fähigkeiten, die er möglicherweise nicht mehr ausreichend entwickelt hat, angewiesen.

Des Weiteren erfüllt der Ort Schule diverse Aufgaben, nicht nur die der Wissensvermittlung. In der Schule lernen Kinder und Jugendliche die sozialen Normen der Gesellschaft, in der sie leben. Sie erarbeiten sich soziale wie kognitive Fähigkeiten, die es ihnen später ermöglichen, sich in ihrem beruflichen und sozialen Umfeld adäquat darzustellen und weiterzuentwickeln. Soziale Kompetenz erwirbt man mit dem menschlichen Gegenüber, das wie ein Spiegel die eigene Persönlichkeit zurückprojiziert. Diesem Ziel des Erwerbs sozialer Kompetenzen dient

die Schule als Ganzes, aber auch diverse unterrichtliche Methoden wie Gruppen- oder Partnerarbeit, gruppenteiliges Arbeiten im Team, Schülerkorrektur und -teaching und vieles mehr. Der vermehrte Einsatz von Chatbots könnte diesen Austausch, das soziale und emotionale Lernen stark in den Hintergrund drängen, Jugendliche werden schon in der Pubertät zu Einzelkämpfern und lernen Kooperation und Kommunikation nur noch am Rande.

ChatGPT kann nicht auf jede Frage eine befriedigende Antwort geben, vor allem wenn Schüler komplexe Fragestellungen auf einfache Faktoren reduzieren. Dass Kontextinformationen, die für eine zufriedenstellende Antwort nötig wären, fehlen, ist Schülern nicht bewusst und führt leicht zu einer Fehleinschätzung der Antwort von ChatGPT.

Ein heißes Eisen in deutschen Schulen: Hausaufgaben. Zwar wird deren Bedeutung für den Lernerfolg immer wieder kontrovers diskutiert, dennoch war von einer Abschaffung nie die Rede. Das könnte sich jetzt ändern. Denn ChatGPT kann jenseits von praktischen Aufgaben auch Hausaufgaben zuverlässig erledigen, wenn die Prompts gut genug formuliert sind. Aber das erlernen Schüler erfahrungsgemäß sehr schnell. Kritiker runzeln hier die Stirn: Wie wird man künftig Inhalte vertiefen, Kompetenzen durch ihre Anwendung erlernen? Ist das klassische Konzept von Schule mit ChatGPT an seinem Ende? Die Frage muss hier natürlich noch offenbleiben, führt aber direkt zu den vielen positiven Effekten, die ganz allgemein durch den Einsatz von Chatbots in der Schule erzielt werden:

ChatGPT ist ein freundlicher Ersatzlehrer, dem keine Frage zu dumm oder lächerlich ist und der sofort und auf verständliche Art und Weise antwortet. Wenn Schüler konkrete Fragen haben oder auch während des häuslichen Lernprozesses Hilfe bei bestimmten Themen oder für die Prüfungsvorbereitung brauchen, können sie diese über den Textgenerator sofort erhalten und auf diese Weise effizienter und gezielt lernen. Wartezeiten auf Antworten oder unverständliche Antworten gibt es nicht mehr.

Beherrschen Lernende die Kunst des richtigen Promptens, also können sie ihre individuellen Lernbedürfnisse formulieren, dann erhalten sie eine Fülle von auf sie zugeschnittenen Erklärungen, aber auch Übungen und weiterführende Hinweise. ChatGPT stellt sich sprachlich auf den User ein und kann auf jedem Niveau und in jedem Stil antworten. Übungsmaterial oder Lückentexte, aber auch Textaufgaben kann der Bot auf Anweisung hin generieren und der Schüler kann seine Lösung mit der des Bots abgleichen. Schüler können Texte oder Audiodateien eingeben und ChatGPT auffordern, ihnen die Sachverhalte vereinfacht darzulegen. Was für eine Möglichkeit!

Vor allem Jugendliche mit besonderen Anforderungen haben mit

ChatGPT ein Werkzeug an der Hand, das immer auf sie eingeht, ohne sie zu deklassieren. Wer sich nicht traut, im Klassenverband eine Frage zu stellen, egal, ob man eine Matheschwäche, Hörprobleme oder Migrationshintergrund hat, wird diese künftig an den Chatbot richten. Der Chatbot wird jeden schwierigen Text auf einfacherem Niveau anbieten, Übersetzungen oder Sachaufgaben und deren Lösungen liefern.

Man kann die Zeit nicht zurückdrehen, was in diesem Kontext bedeutet, dass man ChatGPT nicht mehr ignorieren und auch nicht in der Schule für nicht existent erklären kann. Im Folgenden sollen deshalb diverse fachgebundene Möglichkeiten, die sich aus dessen Verwendung in der Schule ergeben, umrissen werden.

6.2Wie kann ChatGPT fachspezifisch eingesetzt werden?

Das Fach Deutsch

Die Deutschlehrer sind eine Spezies für sich. Der klassische Deutschlehrer wird sich mit Textgeneratoren sehr schwertun, besteht doch eine seiner Kernaufgaben darin, genau das seinen Schülern beizubringen: Texte zu strukturieren und kohärent auszuformulieren. Das kann jetzt ChatGPT für ihn und jeden Schüler übernehmen. Wie also dieses Werkzeug im Deutschunterricht integrieren?

Die Verwendung von textgenerierenden Tools wie ChatGPT setzt voraus, dass die Schüler nicht nur lesen, sondern auch selbst schreiben können. Ein Schüler, der selbst keine individuelle Schreibkompetenz erworben hat, indem er bereits erlernt hat, aktiv Texte zu produzieren, kann von der KI erstellte Texte nicht wirklich beurteilen. ChatGPT kann in bestimmten Situationen das Schreiben bereichern, es auch durch Ideen erleichtern, aber im schulischen Rahmen sollten Textgeneratoren den persönlichen Schreibprozess niemals abnehmen bzw. ersetzen, sondern allenfalls ergänzen. Schließlich ist Schreiben auch Ausdruck der eigenen Persönlichkeit, einen Schreibstil zu entwickeln eine höchst individuelle Angelegenheit. Ein Text von ChatGPT ist nie mein Text, sondern der Text einer Maschine, der meiner Persönlichkeit nicht entspricht.

Hinzu kommt, dass man während des persönlichen Schreibprozesses neue Einsichten und ein erweitertes Verständnis über den Gegenstand gewinnt. Schreiben regt kognitive Prozesse an, die wiederum in das Schreibprodukt einfließen. Diese Erfahrung vermittelt dem Schreibenden eine Selbstkompetenz, die ihn selbstbewusster macht und seine Selbstwirksamkeit fördert. Dies bedeutet, dass individuelles Schreiben im

Deutschunterricht nicht ersetzt werden kann.

Dennoch kann man den Einsatz von Chatbots als erweitertes Methoden-Instrumentarium verstehen. Schüler können ChatGPT um Rat fragen, wie sie einen Text stilistisch oder lexikalisch verbessern können, auch Fragen nach der Struktur, der Argumentation oder dem Aufbau eines Textes werden beantwortet und können die Schüler motivieren, an ihrem Text weiterzufeilen. Fehlertexte des Chatbots oder Texte auf einfachstem Niveau schulen die sprachliche Sensibilität und dienen der Erweiterung der Sprachkompetenz, wenn man sie von den Schülern korrigieren lässt. Umgekehrt können Schüler mit Migrationshintergrund ihre eigenen Texte verbessern und sich auch die Fehler erläutern lassen. Lückentexte sind ein probates Mittel, um bestimmte Wortarten zu bestimmen bzw. zu isolieren; diese kann man ChatGPT formulieren lassen und in einem zweiten Schritt auffordern, Synonyme zu finden.

Textproduktion zu einem bestimmten Thema als häusliche Arbeit aufzugeben (zum Beispiel: Schreibe eine Erörterung zu den Vor- und Nachteilen von Taschengeld), hat sich auf klassische Art durch ChatGPT allerdings erübrigt. Möglich wäre hier, die Aufgabe bewusst unter Einbezug von ChatGPT zu formulieren (»Lass ChatGPT einen Aufsatz zu den Vor- und Nachteilen von Taschengeld schreiben«) und den Text in der Schule auf seine Argumentation hin untersuchen zu lassen, was den Schülern sicherlich viel mehr Spaß macht. Weiteres Lernziel besteht hier darin, die Jugendlichen mit den unterschiedlichen Ergebnissen zu konfrontieren und die sprachlichen Eigenheiten zu vergleichen.

Herausfordernd ist die Arbeit mit älteren Schülern. Es empfiehlt sich auf gar keinen Fall, die Verwendung von ChatGPT bei Referaten oder Hausarbeiten rigoros zu verbieten. Dies würden viele Schüler trotzdem tun und den verwendeten Chatbot kreativ verschleiern. Offene Kommunikation und klare Abmachungen sind deshalb wichtig und sollten dazu führen, dass ChatGPT nur im vereinbarten Rahmen verwendet und alle Prompts aufgelistet werden; andere, digitale wie analoge Quellen, sollten unbedingt vereinbart und eingefordert, schließlich auch von der Lehrkraft kontrolliert werden. Faktencheck als Aufgabe der Referatsvorbereitung muss vorher besprochen sein. So lernen die Schüler, ChatGPT als Hilfsmittel zu verwenden, das nicht mit einem Lexikon verwechselt werden darf (man denke an den Common Crawl). Der Vorschlag für die Deutschlehrer könnte deshalb lauten: Im Vordergrund der Bewertung sollten künftig nicht mehr der Inhalt und Aufbau des Referats, sondern der freie und verständliche Vortrag sowie die Beantwortung von Fragen danach stehen.

Fremdsprachen

Für Fremdsprachenlehrende und -lernende ist ChatGPT ein wahrer Segen, weil Erstere darin eine große Arbeitserleichterung sehen, Zweitere einen Tutor zur Seite haben, der in der fremden Sprache absolut sicher ist.

Der Unterrichtende kann endlich individuell fördern, denn ChatGPT kann zum gleichen Thema Arbeitsblätter auf unterschiedlichen Niveaustufen ausgeben, welche von den Jugendlichen mithilfe von ChatGPT oder auch im Team korrigiert werden. Andere Möglichkeit: Texte aus der Fremdsprache werden – als Fehlertexte – vorgelegt und von den Jugendlichen korrigiert. Prüfungsvorlagen oder einfache Abfragen zu bestimmten Grammatikthemen kann man von ChatGPT entwerfen lassen, samt Lösungen natürlich.

Zu Hause können die Schüler dem Chatbot ihre selbst verfassten Texte zum Korrigieren geben und bekommen sogar noch Erklärungen zu ihren Fehlern. Auch Übungsfragen für mündliche Prüfungen kann man ChatGPT stellen, und er unterhält sich wie ein menschlicher Partner in fehlerfreier Fremdsprache mit dem Fragenden. So können die kommunikativen Fähigkeiten verbessert und das Selbstbewusstsein gestärkt werden. Individuelle Fragen, die man sich im Unterricht nicht auszusprechen traut, können dem Chatbot gestellt, Aussprachehilfen können erfragt werden, der Entwurf einer Übungsprüfung, Übersetzungen von Passagen oder einzelnen Wörtern, Wortschatzerweiterungen, Interpretationen von Gedichten und kreative Ideen, um einen Text in der Fremdsprache zu erstellen – alles ist möglich.

Wichtig erscheint auch hier, klare Absprachen zu treffen und den Schülern klarzumachen, dass das Ziel darin besteht, die Sprache selbst zu beherrschen. Dies klappt natürlich nur, wenn man Vokabeln lernt und die Sprachkompetenz aktiv erwirbt. Dennoch: Mit dem Chatbot kann man sich schon mit wenigen Grundkenntnissen in der fremden Sprache wie mit einem lebendigen Gegenüber austauschen und so seine Fertigkeiten enorm erweitern.

Naturwissenschaften und Mathematik

ChatGPT ist ein wahrer Fundus an Wissen zu allen Themen, die einen je interessiert haben oder noch interessieren könnten. Die Schüler können hier wirklich nach Herzenslust recherchieren und Themen in der Tiefe erkunden. Egal, ob das Liebesleben der Schimpansen, die

Relativitätstheorie, die Mitternachtsformel oder das Kreidezeitalter von Relevanz ist, ChatGPT wird darauf in der gewünschten Form antworten. Dies bedeutet, dass Schüler durch den Prompt »Bitte in einfacher, verständlicher Sprache« Inhalte erfragen können, die sie in der Schule noch nicht verstanden haben oder selbst gern weiter vertiefen möchten. Um Stoff zu wiederholen, kann man ChatGPT auffordern, Fragen zu einem bestimmten Thema zu stellen oder einen Lückentext zu entwerfen. Auch wer Multiple Choice bevorzugt, dem sind keine Grenzen gesetzt:

```
>>Thema Schimpansen. Erstelle ein Multiple Choice zu den 20
wichtigsten Fakten, Niveau: 8. Klasse Gymnasium
```

Umgekehrt kann man Chat GPT auch dialogisch bitten, das eigene Wissen abzufragen und so Lücken herausfinden:

```
>>Thema Plattentektonik: Frage mich nach den Auswirkungen der
Bewegungen der Kontinentalplatten
```

Der Schüler lernt aktiv, bekommt auf der einen Seite Bestätigung und kann andererseits noch nicht erlernten Stoff verfestigen. Vorteil: Er oder sie wird von einem neutralen und immer dialogbereiten Lernbegleiter unterstützt, dem nie der Geduldsfaden reißt.

Im Fach Mathematik ist GPT-4 bereits auf einem besseren Weg als die Vorversion, scheint aber noch immer kein zuverlässiger Lernbegleiter zu sein. Grafische Darstellungen bereiten Schwierigkeiten und werden mit einem erklärenden Text ersetzt, Rechnungsergebnisse sind zum Zeitpunkt der Erstellung des vorliegenden Buches noch nicht verlässlich.

Deutsch als Fremdsprache oder Zweitsprache

Jugendliche mit Migrationshintergrund könnten zu den größten Gewinnern von textgenerierenden Assistenten werden, wenn sie es verstehen, diese in ihrem Sinne zu bedienen. ChatGPT bietet eine Vielzahl an Möglichkeiten, natürlich angefangen bei der sogenannten Kompetenzunterstützung, die darin besteht, dass ChatGPT mit dem Schüler Gespräche führt, in denen dieser auf seine Fehler aufmerksam gemacht wird und ihm diese anhand einfacher Sprache erklärt werden. Das Erläutern von Vokabeln und Grammatik, aber auch das Erklären von rätselhaften Riten (»Warum wirft die Braut ihren Brautstrauß in die Höhe?«) oder auch Alltagssituationen (»Wie verhält man sich in einer deutschen Kirche?«) bereitet ChatGPT keine Probleme und kann kulturelle Barrieren abbauen.

Formulierungen in der eigenen Sprache kann ChatGPT natürlich fehlerfrei ins Deutsche oder in eine der anderen (insgesamt 95) Sprachen übersetzen, was vor allem bei Bewerbungen oder der Vorbereitung wichtiger Gespräche oder Prüfungen vorteilhaft ist. Außerdem kennt ChatGPT alle Lernhelfer wie Bücher, Apps (Duolingo, Rosetta Stone u. a.), Websites (Tandem, Busuu, Deutsche Welle) oder Videos, die das Lernen der neuen Sprache und Kultur erleichtern. Fachspezifische Inhalte zum Beispiel aus dem Mathe- oder Geschichtsunterricht kann ChatGPT in einfacher Sprache erläutern; auch digitale Texte aus dem Unterricht, die in den Prompt eingegeben werden, kann sich der Schüler in einfacherer Sprache ausgeben lassen.

Im Endeffekt können zwei wesentliche Dinge gleichzeitig stattfinden: Wer Deutsch als Zweitsprache erlernt, den unterstützt ChatGPT dabei, ohne dass der Schüler die eigene Muttersprache vergisst. Das trägt zum Erhalt der eigenen kulturellen Identität bei und öffnet den Zugang zu einer anderen. Im besten Fall wird in beiden Sprachen auf höherem Niveau kommuniziert.

6.3Absolut unerlässlich: Der Faktencheck

Im Zeitalter von Fake News ist eine der Kernkompetenzen, die in der Schule gelehrt werden müssen, der Faktencheck. Mit ChatGPT ist diese Notwendigkeit nicht obsolet geworden, ganz im Gegenteil. Wer seine Schüler dieses Sprachmodell verwenden lässt, ist dazu verpflichtet, sozusagen parallel den Faktencheck als unerlässlichen Arbeitsschritt einzuführen und konsequent einzufordern. Große Teile des Textvolumens entstammen dem Common Crawl, was allein ausreicht, um so manche Aussage von ChatGPT zu hinterfragen. Leider füllt ChatGPT eigene Lücken durch kreative Erfindungen bzw. ideenreiche Kombination von Informationen, sodass man vor allem dann, wenn man sich nicht auskennt, in die Falle tappen kann.

Da hilft nur eins: Fakten überprüfen. Der Faktencheck ist die basale Kernkompetenz für unser künftiges Leben, das Bewerten von Aussagen auch im politischen Kontext. Wichtig ist diese Überprüfung auch deshalb, weil viele Gymnasiasten nach der Schule akademische Ausbildungswege anstreben, und diese sind nicht mit ChatGPT (allein) zu haben. Lesen und Auswerten von Quellen sollte also wie in den Zeiten vor KI einen definierten und klaren Stellenwert im Unterricht haben. Schülern muss vermittelt werden, am besten anhand von Beispielen, dass ChatGPT völlig unbemerkt Fakten schafft, auch wenn diese keine Übereinstimmung mit der Wirklichkeit haben.

Wenn man, um einen kleinen Test durchzuführen, ChatGPT nach Schillers »Schöner Seele« fragt, bekommt man zwar eine weitgehend zutreffende Antwort, aber leider eben nicht in jeder Hinsicht: Schiller hat nämlich nie einen Briefroman – so ChatGPT in einer Antwort – verfasst, in dem er sein Konzept der Schönen Seele erläutert. Überhaupt: Einen Briefroman in diesem Zeitraum gibt es nur von Goethe. Hier scheint etwas miteinander verwurstet worden zu sein …

6.4Auf lange Sicht: Was ändert ChatGPT in der Schule?

Unterricht und die Rolle des Lehrenden

Der Lehrende bleibt ganz am Ende immer auch Prüfender, aber der Weg dorthin wird sich durch Digitalisierung, KI und nun ChatGPT stark verändern und zu völlig neuen Unterrichtsformen führen. Lehrkräfte stehen künftig insgesamt weniger im Zentrum, da sie zwar Inputs geben, aber die Schüler mit konkreten Aufgaben in die Arbeit mit Chatbots schicken. Das stärkt die Schüler dabei, einerseits selbstständiger zu werden und andererseits mehr zu kooperieren. Die Lehrkraft kann ChatGPT gezielt in Gruppenaktivitäten einbinden, wenn es darum geht, gemeinsam Lösungen zu erarbeiten.

Im Vordergrund wird in Zukunft also nicht mehr die Rezeption des Stoffs, der durch die Lehrkraft vermittelt wird, stehen, sondern die Analyse, Überprüfung und Bewertung von Inhalten der Textgeneratoren. Diese können bei einem kontroversen Thema sogar als Dialogpartner dienen und die Schüler darin schulen, Gegenpositionen zu formulieren. Speziell in den Sprachen und Gesellschaftswissenschaften lässt sich ChatGPT deshalb sehr zielorientiert einsetzen.

Die Befürchtung, Schüler würden dann nur noch abschreiben, lässt sich nicht halten, schon allein deshalb, weil alle Schüler für den identischen Prompt eine leicht abgewandelte Antwort erhalten. Die Schüler müssen sich damit kritisch auseinandersetzen und noch andere Quellen befragen, was man wiederum in Form von Partnerarbeit organisieren kann. Der Unterricht wandelt sich auf diese Weise stark weg vom passiven Rezipieren in Richtung aktiver Teilnahme und Gestaltung des eigenen Lernprozesses. Schließlich werden damit sogar die 4Ks des Lernens im 21. Jahrhundert erfüllt (Kritisches Denken, Kollaboration, Kreativität, Kommunikation).

Die Rolle der Lehrkraft wird sich ändern. Bei der didaktischen Aufbereitung von Unterrichtsstoff werden die Lehrkräfte Textgeneratoren wie ein Tool integrieren und dadurch das Repertoire an

Methoden stark in Richtung Eigenständigkeit der Schüler sowie Medienkompetenz erweitern. Dadurch fungieren Lehrkräfte zunehmend als individuelle Lernbegleiter, die während des Unterrichts kaum mehr vor den Schülern an der Tafel stehen, sondern sich als Berater zwischen den Lerngruppen bewegen, Tipps geben, neue Lernaufgaben oder Arbeitsblätter für Einzelne generieren und insgesamt eher als Partner wahrgenommen werden.

Bewertung von häuslichen Arbeiten

Beim Thema Bewertung von häuslich erstellten Leistungen wie Referaten oder Hausarbeiten wird es zunächst Probleme geben. Schüler neigen dazu, den einfachsten Weg, sprich den Weg des geringsten Widerstands zu gehen. Deshalb muss klar geregelt werden, welche Rolle ChatGPT spielen darf und in welchem Umfang es im Vergleich zu gängigen Quellen eingesetzt wird. Den Schülern muss klar sein, dass sie sich nicht auf die Ergebnisse von ChatGPT verlassen dürfen; sie im Vorfeld zu einem Abgleich der Chatbot-Information mit anderen Quellen anzuregen oder sogar aufzufordern, erscheint akademisch und pädagogisch betrachtet sinnvoll. Außerdem ist denkbar, dass der Schüler den verwendeten Prompt samt Antwort des Textgenerators in die Quellenangaben einarbeiten muss, was bereits vom Schulministerium Nordrhein-Westfalen in einer Handreichung empfohlen wird. Daran lässt sich sowohl ablesen, wie umfangreich der Einbezug der KI war, als auch, mit welcher Kompetenz der Schüler mit dem Chatbot gearbeitet hat.

Eine Bewertungsmöglichkeit: Je höher der Anteil von ChatGPT an der zu erbringenden Leistung, desto höher die Anforderungen an die zu erbringende Leistung. Ein weiterer Ansatzpunkt wäre, die Präsentation des Referats, die Kommunikation mit den Zuhörern und Aspekte wie Lautstärke, Verständlichkeit, Körperhaltung einzubeziehen.

All dies sollte im Fachkollegium, noch besser im Rahmen der Schule bzw. einer Konferenz, diskutiert, geklärt und beschlossen werden. Je klarer die Regelungen für den Umgang mit ChatGPT, desto weniger Stolperfallen und Schwierigkeiten warten auf alle Beteiligten.

Dazu gehört auch eine wiederholte Einführung in die Arbeitsweise von ChatGPT und die nicht unerheblichen Probleme in Bezug auf das Urheberrecht in jeder Klassenstufe von Neuem. Die Schüler müssen die Chance haben zu begreifen, dass sie es nicht mit einem verlässlichen Nachschlagewerk zu tun haben und deshalb bei allen Arbeitsaufträgen, aber vor allem bei Referaten und Hausarbeiten entsprechend sorgfältig zu Werke gehen müssen.

Entwicklung von Kompetenzen und intellektueller Reife

Richtig eingesetzt, können sprachbasierte KI-Systeme einen entscheidenden Beitrag zur Entwicklung Jugendlicher beitragen. Die oft genannte Medienkompetenz muss hier gar nicht extra erwähnt werden. Viel wichtiger erscheint der Beitrag zur Selbstkompetenz: Gerade Schüler, die durch die intensive Beschäftigung mit den sozialen Medien in einen eher passiven Modus geraten, können im Kontext mit Chatbots wieder aktiviert werden und durch dialogische Aufgaben und anspruchsvolle Arbeitsanweisungen Selbstwirksamkeit erfahren (zum Beispiel: »Vergleiche zum Thema ›Die Auswirkungen des Versailler Vertrages‹ die Ergebnisse von ChatGPT, der Webseite *bpb.de* und *wikipedia.org*«).

Der Anspruch, nicht mehr nur als Konsument am Unterricht teilzuhaben, sondern den eigenen Lernfortschritt aktiv selbst zu gestalten, wird erst allmählich, über die Jahre und eher unbewusst in das Bewusstsein der Schüler einsickern, aber nachhaltig wirken. Versteht es die schulische Didaktik außerdem, in die Arbeit mit Bots sozialfördernde Arbeitsformen zu integrieren, profitieren die Schüler auch in sozialer Hinsicht von ChatGPT. Niemand muss den ganzen Tag in der Schule verbringen, allein in sein Tablet starrend, wie es manch dystopische Ansätze oder Filme suggerieren. Die Schule der Zukunft ist smart und wird kein dissozialer Ort. Wenn elektronische Medien und Sprachassistenten als Tool eingesetzt werden, haben diese sogar eine sehr soziale Komponente: Die Vorschläge von ChatGPT ergänzen und bereichern den Unterricht, wie andere Arbeitsmittel auch.

Im Endeffekt ist durch den Einsatz von dialogischer KI eine Jugend zu erwarten, die besser auf das Leben außerhalb der Schule vorbereitet ist, als dies momentan der Fall ist. Elternverbände und Universitäten monieren immer wieder die unzureichende Tauglichkeit der schulischen Bildung in Bezug auf die Zukunft der Jugend. Dies wird sich jetzt mit großer Wahrscheinlichkeit ändern.

Zu erwartende Probleme und Risiken

Jede Medaille hat zwei Seiten, oder: Wo Licht ist, finden wir auch Schatten.

Thema **Datenschutz**: Die Server stehen in den USA, Schüler und Studenten aus Europa füttern ChatGPT mit sensiblen Daten und bieten damit Einblick in ihre persönlichen Fähigkeiten und Leistungen. Wer schützt diese Daten? Die italienische Datenschutzbehörde hat ChatGPT

im Frühsommer des Jahres 2023 verboten, mit dem Hinweis, dass dieses Programm gegen die Datenschutz-Grundverordnung der EU verstoße. Dem vorausgegangen war eine Datenpanne bei ChatGPT. Einen Monat später gab Sam Altman, CEO von OpenAI, in den Medien bekannt, dass sein Produkt in Italien wieder verfügbar sei. Mit welchen Maßnahmen man die Behörde von einer Aufhebung des Verbots überzeugte, ist in der Öffentlichkeit nicht wirklich bekannt, was zeigt, dass Vorsicht geboten ist.

Thema **Förderung** leistungsschwächerer Schüler: Auf dieser sollte bei der ganzen Diskussion und der Einführung von ChatGPT besonderer Fokus liegen. Bisher hat sich gezeigt, dass bei digitalen Lehr- und Lernangeboten vor allem die leistungsstarken Schülergruppen profitiert haben. Auch bei ChatGPT ist dies nicht ganz auszuschließen, Lehrkräfte haben hier besondere Verantwortung und müssen als Lernbegleiter schwächere Schüler nicht nur mit angepasstem Lernmaterial versorgen, sondern sie auch zu selbstständiger Arbeit mit dem Bot anregen. Gerade diese Schüler müssen lernen, mit diesem Tool so umzugehen, dass es sie in ihrer kognitiven Entwicklung unterstützt und sie nicht von anderen abgehängt werden, aber das kann nur mit angepasster Unterstützung funktionieren. Ansonsten ist zu erwarten, dass die Schere weiter aufgeht und soziale Disbalancen weiter zementiert werden.

Thema **Hausaufgaben**: Zwar wurde dieses Konzept bereits vielerorts und auch sicherlich nicht zu Unrecht in Zweifel gezogen: Kinder sollten in der Schule lernen und Inhalte dort vertiefen, schriftliche Hausaufgaben dürften aus Gründen der Bildungsgerechtigkeit nicht auf den häuslichen Bereich ausgedehnt werden usw. Diese Diskussion ist mit ChatGPT eigentlich vom Tisch. Denn genau dies beherrscht die textgenerative KI: Texte erstellen, egal zu welchem Thema. Für die Hausaufgaben in geisteswissenschaftlichen Fächern, in denen recherchiert, abgewogen, bewertet und eigenständig formuliert werden soll, oder auch in Fremdsprachen stellt ChatGPT oberflächlich betrachtet eine große Erleichterung für die Schüler dar. Theoretisch und praktisch kann ChatGPT die meisten textbasierten Aufgaben perfekt – oder auch mit Fehlern, wenn gewünscht – übernehmen. Da Schüler bekanntermaßen den leichtesten und schnellsten Weg gehen, wird sich die Schule hier möglichst rasch neue Wege einfallen lassen müssen.

Thema **Betrugsversuche**: Es gibt immer noch etliche Leistungsnachweise, die zu Hause vorbereitet und erstellt werden, dazu gehören Referate, Auswertungen, Seminararbeiten und je nach Bundesland etliches mehr. Bis März 2023 musste man wenigstens einige Wikipedia-Artikel lesen, auswerten, neu zusammenstellen, strukturieren und umschreiben. Dies kann sich die Schülerschaft jetzt sparen beziehungsweise elegant umgehen und einen von ChatGPT generierten

Text als eigenen ausgeben. Ein Beispiel soll helfen, diesen Gedanken zu verstehen: Im Deutschunterricht der Oberstufe stehen die sogenannten *Weimarer Klassiker* auf dem Programm. Lehrkräfte geben Referatsthemen aus, um Zeit einzusparen und sich selbst auf die wesentlichen Aspekte konzentrieren zu können. Ein Thema im Kontext Klassiker könnte lauten: »Goethe in Leben und Werk«. Was machen die Schüler seit November 2022? Sie geben einen möglichst umfangreichen Prompt bei ChatGPT ein und erhalten einen zu 98 % tadellosen, sinnvoll strukturierten und nur in wenigen Details zu überarbeitenden Text. Die mit ChatGPT noch nicht vertraute Lehrkraft wird den auf dieser Basis gehaltenen Vortrag mit einer sehr guten, mindestens guten Note bewerten, weil sie davon ausgeht, dass der Schüler sich intensiv mit Quellen, analogen wie digitalen, auseinandergesetzt hat. Dieses Beispiel kann man auf beinahe jedes Schulfach anwenden, und man wird zum immer gleichen Ergebnis kommen: Das bisherige Instrumentarium der Institution Schule, Prüfungsformen eingeschlossen, hat sich spätestens seit der Existenz von ChatGPT erübrigt und bedarf dringend der schnellen und gründlichen Aktualisierung.

Kultusministeriell hat sich noch nichts geregelt, und das ist auch noch nicht so schnell zu erwarten. Die Politik hat es bisher bei Absichtserklärungen bewenden lassen. Jede Schule muss selbst sehen, wie sie hier zurechtkommt und klare Regelungen schaffen, auf deren Einhaltung sie pocht. Und nicht zuletzt müssen die Schüler erkennen, dass sie sich selbst auf die lange Sicht schaden, wenn sie das eigene Denken und Reflektieren an eine KI abgeben und deren Ergebnisse für die eigenen erklären.

An dieser Stelle ist noch zu erwähnen, dass Bildungsinstitutionen an **Technik-Ausfall** gewöhnt und zudem nur selten mit perfekt funktionierender Technik ausgestattet sind. In vielen Schulen gibt es noch kein funktionierendes WLAN, geschweige denn Beamer. Selbst in Großstädten arbeiten viele Lehrkräfte mit Material, das nicht entfernt digitalen Standards entspricht. Andernorts bricht gern das WLAN bei intensiver Nutzung zusammen, oder die smarten Geräte verfügen nicht über die nötigen Programme. Grund: Sie wurden vom Elternbeirat noch nicht abgesegnet, sind zu teuer oder es gibt niemanden, der sie installieren könnte.

6.5Die Vision: Schule der Zukunft

An dieser Stelle sei ein Ausflug in die noch unbekannte Zukunft erlaubt. Visionen beeinflussen die Gestaltung dessen, was vor uns liegt. Warum

also nicht ein bisschen träumen, fantasieren und dabei neue Räume gestalten?

In der Schule der Zukunft hat jeder Schüler in seinem Klassenraum einen Arbeitsplatz, der mit einem smarten Gerät – noch können wir uns nichts anderes als ein Tablet vorstellen – ausgestattet ist. Dieses Gerät ist mit dem schulischen Netzwerk verbunden, auf das alle Arbeiten und Arbeitsblätter geladen werden. Im Arbeitsraum sind 20 weitere Schüler des gleichen Alters, aber nicht unbedingt der gleichen Kompetenzen. Man kennt sich gut, weil man die ersten acht Jahre seines Schullebens in der Schule unweit der eigenen Wohnung gemeinsam verbringt (Sprengelregelung).

Die Schule beginnt um 8 Uhr. Aber nicht mit einem beliebigen Fach oder Unterricht, sondern mit der Klassenleitung, die nach vier Jahren wechselt. Die ersten 15 Minuten werden mit Digital- und Sozialpflege verbracht: Wer hatte Probleme mit dem Hochladen der Arbeitsergebnisse, wer hat Geburtstag, wer kauft für die nächste schulische Aktion ein, wer kontaktiert die kranken Mitschüler. Auch eine kurze Konzentrations- oder Sozialübung ist denkbar.

Dann startet der eigentliche Unterricht, wobei die erste fachliche Einheit des Tages – wie alle anderen – mindestens 100 Minuten umfasst. Einige Schüler verlassen möglicherweise den Klassenraum, weil sie statt Physik das Fach Chemie oder statt Englisch Chinesisch gewählt haben. Unterricht findet in verschiedenen Formen statt, besteht aber zu mindestens 60 % aus Eigenaktivität der Schüler. Der Input durch die Lehrkraft ist nach wie vor unerlässlich, zum Beispiel bei der Einführung in eine komplizierte mathematische Formel oder der Information über den Aufbau einer Textanalyse in sprachlichen Fächern. Nach dem digital gestützten und allen Schülern auf den eigenen Tablets zugänglichen Einführungsvortrag geht es an die Arbeit: Die Schüler erhalten Arbeitsaufträge, die sie auf ihren Geräten abrufen können. Die KI erkennt ziemlich schnell, wo die jeweiligen Probleme eines Schülers liegen, und stellt sofort neue Aufgaben zur Verfügung, meldet dies gleichzeitig der Lehrkraft zurück, welche dadurch Einblick in die Leistungsfähigkeit der Schüler erhält. Sie bewegt sich während dieser Phase aktiv durch den Klassenraum und bietet Unterstützung an. Schüler, welche sich schnell mit dem neuen Stoff vertraut gemacht haben, werden vom Bot erkannt und mit einer anderen Person im Raum vernetzt. So entstehen inhaltlich und sozial neue Arbeitseinheiten, welche sich dialogisch mit neuen Problemen im Umfeld des Stoffs beschäftigen. Projekte oder komplexere Aufgaben für die nachmittäglichen Einheiten werden der Lehrkraft vom Chatbot angeboten, welche diese nach Abwägung entsprechend verteilt. Der Lehrer erkennt außerdem über die Software am erreichten

Bearbeitungsniveau der Arbeitsblätter, wann der Stoff von allen Schülern verstanden wurde und beendet die Einheit beizeiten mit einer alle Schüler integrierenden Idee. Diese ganze Phase kann sich über mehrere Tage bzw. Unterrichtsstunden verteilen.

Der Nachmittag wird mit Vor- und Nachbereitung von Stoff – Vokabeln lernen, Aufsätze schreiben, Berechnungen anstellen –, gemeinsamen Projekten, die sich aus dem vormittäglichen Unterricht ergeben, oder Prüfungen verbracht; die Schüler müssen in diesem Kontext lernen, wie sie diese Zeit effektiv nutzen und können und werden dabei von Chatbots unterstützt. Der Schultag endet zwischen 16 und 16:30 Uhr. Danach fallen keine Arbeiten für die Schule mehr an. Freitags endet der Unterricht nach der ersten Nachmittagseinheit. Für das Wochenende werden im Normalfall keine schulischen Arbeiten angesetzt.

Die Schüler melden sich bei der Lehrkraft selbst zur Prüfung über den jeweiligen Stoff. Prüfungen werden automatisch generiert und finden zweimal pro Woche zentral in einem dafür vorgesehenen Raum statt und werden dort nachmittags von zwei pädagogischen Mitarbeitern durchgeführt.

Schule ist hier ein abgeschlossener Raum, der weniger in das private Leben der Schüler hineinreicht. Die Wochenenden bleiben frei von schulischen Verpflichtungen, dafür umfasst ein Schultag mehr als acht Stunden. Die Schüler lernen einerseits durch den Einsatz von Chatbots effektiv und werden ihren eigenen Möglichkeiten entsprechend gefördert. Andererseits gestalten sie ihren eigenen Arbeitsprozess bis hin zum Zeitpunkt der Prüfungen. Selektion bzw. Aufteilung auf verschiedene Schularten ist durch die digital erfolgende Differenzierung nicht mehr nötig. Dass innerhalb einer Schule jedoch in den Naturwissenschaften, Mathematik und Sprachen von vornherein unterschiedliche Niveaus angeboten und gewählt werden können, wäre wünschenswert.

Die Lehrkräfte übrigens sind – wie in Finnland – die jeweils Besten ihres Abschlussjahrgangs und dürfen nur nach einer Qualifizierung für den Lehrberuf (Form: Auswahlgespräch, schriftliche Bewerbung) das Studium des Lehramtes aufnehmen. Ihr gesellschaftliches Ansehen ist deshalb sehr hoch, ihre Bezahlung entsprechend.

Selbstredend knüpft diese Vision an bestehende Modelle an, aber sie geht doch durch den konsequenten Einsatz von KI weit darüber hinaus. Was hoffentlich auch sichtbar wird: Lehrkräfte werden nicht überflüssig, im Gegenteil, aber ihre Rolle wird völlig neu definiert. Lehrerzentrierter Unterricht gehört der Vergangenheit an, selbst- und KI-gesteuertes Lernen steht im Vordergrund.

6.6KI-Schule – der Pilotversuch

Insgesamt können Chatbots eine wertvolle Ergänzung im Bildungsbereich sein, die Schülern zusätzliche Ressourcen und Unterstützung bietet. Die Integration von KI erfordert jedoch eine sorgfältige Planung, um sicherzustellen, dass ihre Anwendung den pädagogischen Zielen entspricht und die Vorzüge der menschlichen Interaktion nicht beeinträchtigt werden. Noch können wir uns nicht vorstellen, dass und in welchem Umfang sich Schule verändern wird, halten wir doch nur allzu gern an alten Strukturen und Erinnerungen fest.

Dass die KI-Schule jedoch im Bereich des Möglichen liegt, soll ein Modell zeigen, das seit dem Schuljahr 2022/23 durchgeführt wird: Das bayerische Kultusministerium hat an 15 verschiedenen Schulen ein Pilotprojekt gestartet, das in einem Zeitraum von fünf Jahren herausfinden soll, wie KI in der Schule gewinnbringend eingesetzt werden und dabei helfen kann, Schüler unterschiedlichster Altersstufen gezielt zu fördern. Im Zentrum steht dabei die »Erwartung (…), der wachsenden Heterogenität in Lerngruppen viel besser gerecht zu werden«[5]. Die Lehrkräfte wurden und werden fortlaufend gezielt ausgebildet, um die Daten, die durch die KI erhoben werden, auswerten zu können. Was an anderen Schulen noch Zukunftsmusik ist, wird hier bereits Realität, da die Diagnose der Lernprozesse dazu führt, dass jeder Schüler durch auf ihn angepasstes Lernmaterial individuellere Unterstützung beim Lernen erhält.

In den Klassenräumen einer der teilnehmenden Grundschulen, nämlich in Jettingen-Scheppach, befindet sich neben der alten grünen Tafel ein überdimensionaler Flachbildschirm, der mit dem Tablet der Lehrkraft verbunden ist. Die Schüler sitzen über ihre Tablets vertieft und setzen sich mit KI-generierten Inhalten auseinander. Sie geben in der 4. Klasse beispielsweise Daten in das Programm ein, damit dieses im Anschluss daraus einen Bericht erstellt. Die Schüler schreiben diesen Bericht nicht mehr selbst, sondern überprüfen nur das Ergebnis: Hat der Chatbot den Text so verfasst, dass er mit den eingegebenen Daten übereinstimmt?

Dies klingt auf den ersten Blick banal und wird sicherlich von dem einen oder anderen überlesen, aber es stellt eine echte Revolution dar. Denn bisher bestand die entscheidende Kompetenz darin, einen Text selbst zu durchdenken, seinen Aufbau zu strukturieren und dann schließlich in passenden Worten und adäquatem Stil zu Papier zu bringen. Dies übernimmt jetzt die KI bzw. ein Textgenerator. Was hier erprobt wird, ist sicherlich die Zukunft: Texte müssen nicht mehr selbst formuliert, sondern nur mit ihren Eckdaten und Basisinformation in

Programme wie ChatGPT eingegeben werden.

Zurück zur Jettinger Schule: Die Erstklässler üben das richtige Schreiben von Buchstaben auf ihren Bildschirmen. Sie schreiben mit ihren Eingabestiften einzelne Buchstaben und bekommen direkt im Anschluss eine digitale und individualisierte Rückmeldung, ob ihr Buchstabe korrekt ist bzw. ob es noch Nachbesserungsbedarf gibt. Das Programm merkt sogar, ob zu viel oder zu wenig Druck beim Schreiben ausgeübt wird und kommuniziert auch dies. Was der Erwachsene nicht wirklich sehen oder beurteilen kann, übernimmt das Programm.

Alle Schüler wurden mit Tablets ausgestattet, welche die Kinder für die Hausaufgaben auch mit nach Hause nehmen dürfen. Die dort installierten Programme wurden vorher von der Schule aufgespielt. Auf diese Weise kommt KI immer weiter in den schulischen Alltag und wird als Tool genutzt, das zum individualisierten Lernen beiträgt.

Bleibt zu hoffen, dass die Ergebnisse dieses Pilotversuchs auch wirklich Eingang in die Gestaltung der Schule erhalten.

6.7 ChatGPT und die Universitäten

An den Universitäten und Hochschulen herrscht große Unruhe unter den Studierenden und vielleicht noch größere Ratlosigkeit bei den Lehrenden. Eine Bachelorarbeit wird mit 300 Arbeitsstunden veranschlagt und soll den Stand der Forschung in einem bestimmten Bereich abbilden. Der Studierende soll seine Kritikfähigkeit, akademisches Arbeiten und strukturiertes Darstellen unter Beweis stellen und schließt mit dieser Arbeit sein Studium ab. Kann die KI dies künftig besser und schneller? Wie gehen die Universitäten mit dieser Herausforderung um? Vom Bildungsausschuss des Bundestags wurde eine Studie in Auftrag gegeben, die das Problem in seinen Eckpunkten ausloten und Lösungen vorschlagen soll. Man geht zum jetzigen Zeitpunkt davon aus, dass es ohne einheitliche Regelungen wohl nicht gehen wird.

Die Universität Köln hat bereits zum jetzigen Zeitpunkt ein generelles Verbot von ChatGPT für Haus- und Abschlussarbeiten unter Androhung von Sanktionen ausgesprochen, andere Hochschulen verweisen auf die einzelnen Fakultäten, welche dieses Problem eigenständig zu lösen hätten, da die Bedürfnisse sehr unterschiedlich seien. Einzelne Professoren wiederum drohen mit Exmatrikulation, sollte ein Studierender nachweislich sprachbasierte KI verwendet haben. Doch der Nachweis ist problematisch.

Im Endeffekt muss sich auch der universitäre Betrieb den neuen

Realitäten stellen. Sprachbasierte KI-Modelle sind nicht mehr aus unserem Leben wegzudenken und erfordern wohlüberlegte und kreative Vorgehensweisen. Wahrscheinlich wird es dahin münden, dass ChatGPT wie jede andere Quelle künftig gekennzeichnet werden muss. Angeblich arbeiten die Universitäten an dergleichen Lösungen.

Denn auch für Lehrende könnte ChatGPT und dessen Verwendung Vorteile bergen: Administrative Aufgaben, sich wiederholende Prozesse könnten automatisiert werden. Prüfungsfragen kann man nach vielen Jahren endlich variieren, Korrekturen vereinfachen. Warum nicht die sprachbasierte KI nach neuen Hypothesen fragen, Gliederungen umstrukturieren oder Daten analysieren lassen, unbekannte Forschungsergebnisse recherchieren? Natürlich ist der Faktencheck unerlässlich, dennoch werden Lehre und Forschung sich dauerhaft nicht um diese Chancen bringen.

Dann wäre es nicht nur fair, sondern mehr als redlich, auch Studierenden ganz offiziell die Verwendung zuzugestehen. Die Bedenken in Bezug auf Unterschleif, fehlende wissenschaftliche Validität und Datenklau kann man dennoch nachvollziehen.

7 Auswirkungen auf die Arbeitswelt

Bisher wird die Diskussion über KI in den Medien vorwiegend von negativen Schlagzeilen geprägt. Viele Menschen machen sich Sorgen um die Zukunft ihrer Jobs und stellen sich die Frage, ob man sie in fünf Jahren überhaupt noch brauchen wird oder ob die KI ihren Arbeitsplatz längst überflüssig gemacht haben wird. Die Bedenken sind nicht unbegründet, befinden wir uns doch inmitten der vierten industriellen Revolution, die von sprachbasierter KI nun noch zusätzlich befördert wird.

Die britische Telefongesellschaft BT will bis 2030 insgesamt bis zu 10.000 Stellen vor allem in Callcentern durch KI ersetzen lassen, auch der Springer-Verlag kündigte an, im Layout-Bereich künftig KI einzusetzen und damit Stellen zu reduzieren. Die Washington Post will Jobs in den Redaktionen, die sich zum Beispiel mit dem Redigieren von technischen Artikeln oder Produktbeschreibungen beschäftigen, nach und nach durch intelligente Maschinen ersetzen. Das sind alles ernst zu nehmende Meldungen, dennoch muss man medial verbreitete Schlagzeilen immer genauer unter die Lupe nehmen.

7.1 KI in der Arbeitswelt: Welche Jobs sind betroffen?

Die Forschungsabteilung der Investmentbank Goldman Sachs hat sich in einer Studie zur Entwicklung der Arbeitsmärkte geäußert und geht davon aus, dass rund zwei Drittel aller Arbeitsplätze der KI-Automatisierung in irgendeiner Form ausgesetzt sind. Die generative KI könne zu »erheblichen Störungen auf dem Arbeitsmarkt« führen. Langfristig könnten durch generative KI weltweit ungefähr 300 Millionen Vollzeitarbeitsplätze wegfallen. Aber welche Tätigkeiten sind wirklich betroffen von der Disruption, die die KI in der Arbeitswelt hervorbringt? Und um welches Ausmaß an Veränderung handelt es sich eigentlich?

Ein kurzer Blick zurück mag einen Rahmen verschaffen: Die Furcht vor Arbeitslosigkeit durch Innovationen ist nicht neu und zieht sich durch die Geschichte. Die erste industrielle Revolution betraf vor allem Arbeitskräfte, die vorher im ländlichen Bereich handwerkliche, zum Großteil aber bäuerliche Tätigkeiten verrichtet hatten. Sie fanden sich in riesigen Fabriken wieder, in denen sie Maschinen bedienen und in der Massenproduktion arbeiten mussten. Die Mechanisierung führte neben der Steigerung von Effizienz und Produktivität zu inhumanen

Arbeitsbedingungen und zunächst auch zu einer erhöhten Arbeitslosigkeit, da weniger Arbeitskräfte gebraucht wurden, als in die Städte strömten. Aber auf lange Sicht sanken durch die gesteigerte Produktivität die Preise, was zu einer höheren Nachfrage führte und wiederum die Anzahl der benötigten Arbeiter enorm erhöhte. Beispiele für die Möglichkeit gesteigerter Produktivität sind die Erfindung des mechanischen Webstuhls und die Einführung des Fließbands bei der Herstellung von Autos.

Die Disruption der aktuellen, vierten industriellen Revolution bedroht vor allem Tätigkeiten mit starkem Bezug zur Informationsverarbeitung. Plötzlich sind zum ersten Mal die höher qualifizierten Jobs in den Büros, Agenturen und Kanzleien gefährdet. Außerhalb der Gefahrenzone liegen alle nicht digitalisierbaren Berufe: Autos können nicht von der KI repariert werden, elektrische Leitungen wird kein Roboter auf absehbare Zeit verlegen können. KI kann zwar erkennen, wann jemand zu ertrinken droht, aber Bademeister müssen die Rettung übernehmen, das kann kein Roboter. Gleiches gilt für alle klassischen Handwerksberufe vom Dachdecker über den Koch bis zum Heizungsbauer und Installateur. Genau diese Facharbeiter und alle, die in der Pflege oder im sozialen Bereich direkt mit Menschen arbeiten, werden nicht nur jetzt, sondern auch künftig gesucht bleiben.

Sind also die Sorgen um den eigenen Arbeitsplatz begründet? OpenAI-Gründer Sam Altman hatte eigentlich genau das verhindern wollen und seine KI-Forschung ursprünglich als gemeinnützige Stiftung organisiert. KI-Anwendungen sollten nicht dem Profitinteresse einiger weniger dienen und der Gesellschaft insgesamt schaden, sondern mit geeigneten Schutzmaßnahmen in den Händen aller auch der gesamten Gesellschaft nutzen. Die guten Absichten sind im Geschäftsalltag irgendwo auf der Stecke geblieben, inzwischen hat eine gewinnorientierte Tochterfirma einen Deal mit Microsoft geschlossen, das über 10 Milliarden Dollar investiert und dafür alle Lizenzen für die Programme von OpenAI erworben hat.

Zurück zum Arbeitsmarkt: Aktuell betroffen sind also Berufe auf mittlerem und sogar hohem Niveau mit sich wiederholenden, standardisierbaren Abläufen, die vornehmlich am Computer stattfinden: Sachbearbeiter, Rechtsanwaltsgehilfen, Sekretärinnen, Texter, Steuerfachgehilfen, technische Zeichner, Programmierer, Layouter, Datenanalysten – sie werden künftig zwar nicht gänzlich überflüssig, aber ein guter Teil ihrer Arbeit kann von Bots wie ChatGPT übernommen werden. Einfache Anfragen beantworten, kurze Antworten für Mailkontakte generieren, Daten analysieren, Geschäftsprozesse initiieren, ein Produkt zeichnen, einen zielgruppengerechten Werbetext schreiben – das kann ChatGPT entweder ganz selbstständig übernehmen oder

weitgehend vorbereiten. Die Abteilungen, in denen solche Arbeiten typischerweise erledigt wurden, werden mit immer weniger Personal auskommen.

Durch Bots automatisierte Kundenbetreuung anstelle von teuren Callcentern wird sich sicherlich durchsetzen, aber den Kundenservice als solchen wird niemand ganz der KI überlassen wollen. Übersetzer von technischen Produktbeschreibungen werden weit weniger gebraucht, ChatGPT beherrscht zum jetzigen Zeitpunkt schon mehr als 100 Sprachen. Andererseits nimmt die Menge der zu übersetzenden Texte ständig zu, und da die Qualität der maschinellen Übersetzungen noch nicht durchweg hohen Erwartungen entspricht, werden Übersetzer nicht arbeitslos. Das Erstellen von Content für die unterschiedlichsten Zwecke und Zielgruppen kann ChatGPT mit Ideen unterstützen oder gleich komplett texten, also auch Textschaffende und freie Fachautoren haben es unter Umständen schwer, da Firmen möglicherweise auf sie verzichten. Andererseits erleichtert ChatGPT gerade ihnen die Arbeit und macht sie wesentlich produktiver. Für die Arbeitgeber hingegen ist die Nutzung von Chatbots kostengünstig und effektiv, weshalb hier mit dem künftigen und vermehrten Einsatz von KI realistisch zu rechnen ist.

7.2Neue Perspektiven in alten Jobs

Folgt man der Sichtweise von Jens Südekum, Berater der Bundesregierung und einer der einflussreichsten Ökonomen Deutschlands, so wird die KI standardisierte Tätigkeiten übernehmen – auch von Anwälten, Ärzten, Journalisten und Programmierern. Deren Jobs sind noch nicht gefährdet, aber sie werden sich qualitativ verändern. Den langweiligeren Teil der Arbeit erledigt die KI, der Arzt erhält dafür mehr Freiraum, um sich entweder fortzubilden oder mit schwierigen Fällen länger zu beschäftigen. Ähnlich der Journalist, dem die KI das Schreiben von Texten zu einem großen Teil abnehmen kann: Er kann sich endlich intensiv um die Recherche, möglicherweise zu diversen Themen gleichzeitig, kümmern. KI nimmt also Arbeit ab und schafft dadurch Raum für mehr Qualität. Bleibt zu hoffen, dass jemand diese Qualität schätzt und auch bezahlt, sonst geht alles in der Rationalisierung unter.

Gesellschaftliche Schreckensszenarien treffen oft nur in abgemilderter Form ein, meistens deshalb, weil der Faktor Mensch eine Rolle spielt oder aus moralischen Gründen. So wäre es längst möglich, den Beruf der Kassiererin komplett abzuschaffen, den Zugang zu allen Supermärkten per QR-Code zu steuern und diese 24/7 offen zu halten. Dergleichen

Versuche gibt es bereits in etlichen Großstädten Deutschlands, und es werden wöchentlich mehr. Dennoch ist nicht damit zu rechnen, dass in absehbarer Zeit jeder Supermarkt ganz ohne sichtbares Personal betrieben wird, das scheinen die Kunden nicht zu wollen und sie werden es nicht zulassen. Der sogenannte *menschliche Faktor* ist nicht ersetzbar. Zum Glück nicht!

Industrie 4.0 wird auch nicht bedeuten, dass die meisten Berufe, die mit mittlerem Anforderungsniveau am Rechner zu erledigen sind, komplett aussterben. Aber die Berufstätigen werden sich umstellen und daran gewöhnen müssen, dass ein Teil ihrer Arbeit von KI übernommen oder ergänzt wird. Möglicherweise löst diese Entwicklung so ganz nebenbei ein dringendes gesellschaftliches Problem, nämlich den Fachkräftemangel. Wie viele Arztpraxen und Steuerkanzleien suchen momentan händeringend nach Personal und ahnen vielleicht nicht, dass ihr Problem eigentlich schon teilweise lösbar wäre. Auch in der Autoindustrie, die auf Elektroantriebe umstellt, werden Fachkräfte gesucht, die sich digital auskennen oder ausbilden lassen. Wer bisher einen Motor zusammengebaut hat, bekommt die Möglichkeit, sich umschulen zu lassen und wird seinen Arbeitsplatz dann nicht verlieren.

Die konsequente Einführung von KI wird in vielen Bereichen gewinnbringend sein, nicht nur in der Medizin, der Wissenschaft und Forschung, sondern beispielsweise auch in der Landwirtschaft, in deren Digitalisierung sehr viel Geld gesteckt wird und die viel effektiver gestaltet werden kann, wenn genaue Daten über alle Vorgänge, Wachstumsprozesse, Dünger- und Bewässerungsbedarfe sowie Schädlingszahlen vorliegen. Zwar wird dies nur im Zusammenspiel mit weiterer Zentralisierung zu lösen sein, was man kritisch betrachten muss, aber vielleicht lässt sich damit die drohende Nahrungsmittelknappheit in Zeiten von klimawandelbedingten Dürren oder Überschwemmungen besser in Griff bekommen.

7.3Ist die KI ein Jobkiller oder ein Segen auf dem Arbeitsmarkt?

Diese Frage scheint viele umzutreiben, fundierte Antworten sind immer noch Mangelware. Hinzu kommt das demografische Problem der meisten Industrieländer: Eine immer älter werdende Gesellschaft trifft auf immer weniger arbeitsfähige Menschen und einen akuten Mangel an Fachkräften aller Art. Künstliche Intelligenz könnte für viele Bereiche eine Antwort darauf sein, aber niemand scheint letztlich den Überblick zu haben, ob die höhere Produktivität aus der neuen Technik den

demografischen Wandel vollständig kompensieren kann. Und selbst wenn es von den Zahlen her genau aufgehen würde, haben wir dann nicht trotzdem wieder ein Überangebot in bestimmten Jobs und einen Mangel in anderen Berufen? Übrig bleiben am Ende wohl wieder diejenigen Jobs, die eigentlich keiner machen will.

Der Ökonom Bert Rürup behauptet sogar, die KI bringe mehr Jobs, als sie ersetzen würde. Natürlich dürfte ein Teil der Arbeitsplätze, die wir heute kennen, langfristig verschwinden. Mit sehr hoher Wahrscheinlichkeit werden aber auch neue entstehen, vielleicht auch solche, die wir bisher noch gar nicht kennen. Alle technologischen Revolutionen wie die Industrialisierung und Digitalisierung ließen im Vorfeld massive gesamtwirtschaftliche Beschäftigungsverluste erwarten, die aber nicht eingetreten sind. Auch einer Studie der Internationalen Arbeitsorganisation (International Labour Organization, ILO) zufolge wird es eher mehr als weniger Arbeit durch die neue technologische Revolution geben, die KI wird mehr Arbeitsplätze schaffen als vernichten. Zum gleichen Ergebnis kommt eine groß angelegte IBM-Studie, die zusätzlich auch positive Auswirkungen auf die Qualität der Arbeitsplätze erwartet.

Aber wie sehen diese neuen KI-Jobs wirklich aus? Werden wir als Sklaven auf der *Machine-Learning-Galeere* enden und Daten labeln wie die unterprivilegierten Clickworker in Kenia? Die Netzpiloten[6] nennen konkret vier neue KI-Berufe, die es bisher nicht gab: Der **KI-Trainer** ist als Top-Experte mit der Entwicklung und dem Training von KI-Modellen beschäftigt. **KI-Ethiker** sind mit den ethischen Aspekten der Entwicklung und des Einsatzes von künstlicher Intelligenz befasst und bewerten die Auswirkungen der Technologien auf die Gesellschaft und das Individuum. Noch stärker spezialisiert sind **KI-Sicherheitsexperten**, die KI-Systeme vor potenziellen Bedrohungen und Angriffen schützen müssen. Last but not least wird das Tätigkeitsfeld des **KI-Beraters** skizziert, der Unternehmen und Organisationen bei der Integration und Nutzung von KI-Technologien unterstützt. War für Sie der passende Job mit dabei?

Vergleicht man diese Liste der neuen Berufe mit denen, die künftig wohl wegfallen werden, sieht es dann nicht doch eher nach einem massiven Jobverlust durch die neuen Technologien aus? Sind die Thesen von Bert Rürup und anderen vielleicht nur eine gut gemeinte Beschwichtigung? Opium für das Volk? Zum jetzigen Zeitpunkt scheint das niemand seriös analysiert und umfassend durchgerechnet zu haben. Bis belastbare Aussagen und Zahlen vorliegen, hoffen wir einfach gemeinsam, dass die neue Produktivität tatsächlich den künftigen Mangel an Arbeitskräften nur kompensiert, anstatt gleich eine neue Massenarbeitslosigkeit auszulösen.

7.4 Nicht alle Veränderungen werden positiv sein

Problematisch erscheint der Einsatz von KI dort, wo damit Kontrolle der Produktivität und des Einzelnen einhergeht. Dies könnte sowohl hoch entwickelte als auch ganz einfache Jobs betreffen. Wer einen Computer am Handgelenk trägt, Sensoren am Puls hat oder per GPS ständig mit dem firmeninternen Netzwerk verbunden ist, ist ein offenes Buch und kann permanent überwacht werden. Im Bereich der Logistik, bei Firmen wie Amazon, wird bereits seit Längerem den Pickern der Weg zum Produkt vorgegeben. Darüber hinaus wird vergleichend eingeblendet, welche Schlagzahl die Kollegen haben. Dies erhöht den Stress, aber sicherlich nicht dauerhaft die Leistungsbereitschaft und erinnert unangenehm an die erste industrielle Revolution, als das Fließband eingeführt wurde.

Vielleicht kann die durch KI bedingte Disruption in der Arbeitswelt noch verhindert werden, oder wir befinden uns mitten in einem schleichenden Prozess des Übergangs. Dies kann man zum jetzigen Zeitpunkt nicht abschließend bewerten oder vorhersagen. Dennoch dürfen Arbeitnehmer nicht ihre Augen vor der neuen Realität verschließen und müssen Flexibilität üben. Je zwanghafter und ängstlicher sie an den gewohnten Arbeitsprozessen festhalten, desto stärker wird sie die Umstellung treffen. Möglicherweise wird dann eine Kollegin bevorzugt, wenn es darum geht, Stellen neu zu verteilen. Auch der Bademeister muss sich mit der gefahrenerkennenden KI vertraut machen, genauso wie die Sekretärin, die Standardroutinen an ChatGPT abgibt und dafür mehr Zeit gewinnt für anspruchsvollere Tätigkeiten oder Kundenbetreuung. Ebenso die Jura-Absolventen, sie müssen sich damit genauso auseinandersetzen wie die angehenden Mediziner und Lehramtskandidaten: Wie wird mein Arbeitsfeld in Zukunft aussehen, wo muss und wie kann ich mich anpassen? Und wie kann ich meine eigenen Lebensvorstellungen damit verknüpfen?

7.5 Neue Chancen zur Humanisierung der Arbeitswelt

Andererseits besteht die realistische Möglichkeit, dass die KI, sinnvoll eingesetzt, zur Humanisierung der Arbeit beiträgt, was anhand einiger Aspekte angedacht werden soll:

- **Effizienz**. Mit KI-gesteuerten Tools können Mitarbeiter ihre Zeit effizienter nutzen, zum Beispiel beim automatischen Planen von Meetings oder der automatischen Verwaltung der E-Mail-Kommunikation. Dokumente kann man in Sekundenschnelle nach

einer relevanten Information durchsuchen lassen, statt dies selbst in stundenlanger, ermüdender Arbeit zu übernehmen. Und man bekommt weniger Spam in die Mailbox.

■**Gesundheitsvorsorge**. Arbeitgeber können den Mitarbeitern Smartwatches oder Fitnesstracker anbieten, damit diese während der Arbeit ihre Körperdaten kontrollieren können. Das Ziel wäre die Vermeidung von Stress oder Burn-out. Sprachsensible Apps können den Stimmenton, die Wortwahl und andere Kommunikationsmuster eines Benutzers überwachen und Anzeichen von Stress, Depression oder andere psychische Gesundheitsprobleme erkennen, um dann sofort proaktive Maßnahmen vorzuschlagen. Natürlich ohne dass der Chef dies erfährt, es geht um Unterstützung, nicht um Überwachung.

■**Fortbildung**. Weiterbildung kann stärker individualisiert werden, indem personalisierte Lernpfade für Mitarbeiter erstellt werden, die tatsächlich auf deren Fähigkeiten und Interessen sowie Karrierezielen aufbauen.

■**Ernährung**. Viele Arbeitnehmer klagen darüber, dass die Ernährung durch die Arbeit beeinträchtigt und vor allem *Falsches* gegessen wird. Dem könnten KI-gestützte Apps entgegenwirken, indem sie während der Arbeitszeit die Nahrungsaufnahme verfolgen und Tipps geben, wie diese gesundheitsfördernd optimiert werden kann. Statt den zweiten Energieriegel könnte eine App den Griff zur Banane vorschlagen.

■**Überwachung der Umwelt**. In Großraumbüros klagen Mitarbeitende häufig über Kopfschmerzen und schnelle Ermüdung. Dafür gibt es Abhilfe: Sensoren, kombiniert mit KI, können die Qualität der Innenraumluft, die individuelle Anfälligkeit gegenüber Allergenen oder Schadstoffen und andere umweltbedingte Gesundheitsfaktoren überwachen und wenn nötig Alarm schlagen.

Ranga Yogeshwar, namhafter Wissenschaftsjournalist und Physiker, verweist darauf, dass sich der Begriff *Arbeit* erst im Laufe der ersten industriellen Revolution eingebürgert hat und wir jetzt die Chance haben, aus dieser *Diktatur* auszubrechen, indem wir die hierarchisch geordnete Arbeitswelt verlassen. Verglichen mit den letzten 150 Jahren, ist die wöchentliche Arbeitszeit stark gesunken, und der Trend geht immer weiter.

Darüber hinaus gibt es bereits heute viele Branchen, die ihren Mitarbeitern große Freiräume zugestehen: »Arbeite, wo und wann du willst!« Damit einher geht die sich vor allem in der jungen Generation durchsetzende Haltung, dass Arbeit nur ein Teil des Lebens ist, aber nicht mehr die Selbstdefinition eines Menschen ausmacht. Stichwort: Work-Life-Balance. Hier wird eine große Rolle spielen, dass sich die Akzeptanz von Homeoffice und Remote-Arbeit beschleunigt. Diese

Formen der Arbeit bieten große Flexibilität, ermöglichen einen einfachen Austausch auf internationaler Ebene und schaffen sogar Zugang zu einem globalen Talentpool. Formen wie Freelancing, Projektarbeit und flexible Arbeitszeitmodelle werden viele Berufssparten in noch viel umfassenderer Weise prägen. Kühne Vordenker verknüpfen Arbeit schon mit Urlaub und haben dafür den Begriff *Workation* geprägt. Auf der anderen Seite kann dieser Trend zu immer mehr Freiheit und Eigenständigkeit in Bezug auf die soziale Absicherung mit großer Verunsicherung verbunden sein. Der Einzelne wird aber nicht nur in dieser Hinsicht stärker gefordert sein, auch die lebenslange Bereitschaft zur persönlichen Weiterbildung wird eine künftige Herausforderung darstellen.

Letztlich liegt die Zukunft nicht offen vor uns. Dass KI, Sprachmodelle wie ChatGPT und weitere Automatisierung dominieren und die Arbeitswelt entsprechend beeinflussen werden, daran besteht wohl kein Zweifel mehr.

Abschließend noch ein ganz praktischer Tipp, wenn Sie gerade einen neuen Job finden müssen und Stellenangebote wälzen: Nutzen Sie die BERT, um mit fast 90 %iger Sicherheit Ihr Gehalt nur anhand der Stellenausschreibung vorhersagen zu können. BERT ist ein Sprachmodell, das von Google entwickelt und 2018 vorgestellt wurde. Der Algorithmus analysiert und verarbeitet Texte und lernt daraus, wie Menschen miteinander kommunizieren. Eine Forscherin der Uni Stanford fütterte BERT mit 800.000 Stellenanzeigen und den dazugehörigen Gehältern, um die Zusammenhänge zwischen Ausschreibung und Vergütung zu finden. Nach dem Training musste die KI bei den übrigen 200.000 Stellen voraussagen, welches Gehalt dahinter winkt. Und BERT lag tatsächlich bei 87 % der Fälle goldrichtig, in den restlichen 13 % war es wohl auch nicht allzu weit von der Realität entfernt. Sie sehen: Wer als Arbeitnehmer die Möglichkeiten der KI konsequent nutzt, kann daraus einen echten Vorteil ziehen. Und im nächsten Jobinterview bei den Gehaltsverhandlungen noch eins obendrauf legen.

8 Auswirkungen auf die Gesellschaft

Noch haben wir wohl nicht vollständig begriffen, was eigentlich passiert ist. ChatGPT ist seit über einem Jahr frei verfügbar und stellt eine *disruptive Technologie* dar, die von vielen bereits intensiv genutzt wird. Auch Bildgeneratoren wie Midjourney, Dall·E 2 und Artsmart.ai werden immer stärker genutzt und überfluten uns mit einer ganz neuen Bilderwelt. Wahrscheinlich verstehen wir immer noch nicht komplett, dass wir mit Programmen arbeiten, die ungeahnte Auswirkungen nicht nur auf die Arbeitswelt, sondern auch auf unser Leben und unsere Gesellschaft haben werden. Im Folgenden wollen wir einige der möglichen Veränderungen und neuen Phänomene beschreiben.

8.1 Fake News und Deepfakes

OpenAI bekräftigt, dass es beständig die Risiken der Fehlinformationsverbreitung durch seine Modelle minimiert. Dennoch ist durch sein Sprachmodell die Verbreitung von Fake News (falsche oder irreführende Informationen, die als wahr dargestellt werden) immer noch möglich, wenn diese von ChatGPT nicht als solche erkannt werden. Extreme politische Gruppierungen können dies für ihre eigenen Ziele nutzen und auf diese Weise extremistische Meinungen gezielt und massenhaft verbreiten, indem sie Modelle mit entsprechenden Falschinformationen füttern und die Filter geschickt umgehen. Wahlen können dadurch beeinflusst werden, wie bereits bei den US-Wahlen passiert, oder öffentliche Personen diskreditiert. Je mehr Menschen ChatGPT als alleinigen Informationskanal oder wie ein Nachschlagewerk benutzen, desto größer die Gefahr, dass es manipuliert wird und die User Fehlinformationen nicht als solche erkennen. Fatal, wenn ihre Entscheidungen dann bei einer demokratischen Wahl das Ergebnis von Fake News sind.

Das jüngste Beispiel dazu stammt aus der Slowakei. Hier hat soeben Ex-Premier Robert Fico die Wahlen gewonnen mit mutmaßlichen Folgen für die Ukraine und einem möglichen Richtungswechsel des kleinen EU-Landes von West nach Ost. Die Linksnationalisten könnten zum nächsten Problemfall innerhalb der Europäischen Union werden. Russlands Trollfabriken haben wohl auch hier wieder kräftig mitgeholfen, um unseren Informationsraum mit Halbwahrheiten und Lügen zu verunreinigen. Sie arbeiten am falschen Bild, dass Demokratie auch nicht

besser sei als Autokratie. Und trotz des Verhaltenskodex haben soziale Netzwerke Fake News und Propaganda aus Russland verbreitet.

Erst kürzlich wurde ein detaillierter Bericht der EU-Kommission präsentiert, der den Umgang von Facebook, YouTube und TikTok mit Fake News und politischen Unwahrheiten beleuchtet. Die freiwillige Desinformationscharta der Europäischen Union und der »Verhaltenskodex gegen Desinformation« stammen aus dem Juni 2022. Der Digital Services Act (DSA) wurde mittlerweile als Gesetz eingeführt und verpflichtet große Plattformen, die Risiken von Desinformationskampagnen drastisch zu reduzieren. YouTube hat über 400 Kanäle geschlossen, TikTok entfernte über 3.000 Fake-Accounts aus Russland, bei X (vormals Twitter) ist allerdings noch keine Besserung in Sicht und hier gab es den größten Anteil an Desinformationsbeiträgen. Was wir daraus lernen können: Die Flut von Fake News und Desinformation ist so mächtig, dass unsere Dämme nicht standhalten. Und nach den Wahlen gelten die neuen Machthaber trotzdem als »demokratisch legitimiert«.

Neben ChatGPT als Textgenerator gibt es, auch von OpenAI entwickelt, den Bildgenerator DALL·E. Dieser wandelt die per Prompt sprachlich erfolgte Beschreibung eines gewünschten Bildes in ein solches um. Die Ergebnisse sind frappierend echt! Aber auch hier besteht ein großes gesellschaftliches Risiko, nämlich das der Täuschung oder Manipulation durch einen Deepfake. Dieser ist ein mittels KI erzeugtes Bild, eine Audioaufnahme oder eine kurze Filmsequenz, die so realistisch wirken, dass es praktisch unmöglich ist, sie als Täuschung (Fake) zu identifizieren. In diesen Deepfakes werden Situationen vorgetäuscht, die niemals stattgefunden haben. Sie können deshalb für gezielte und absichtliche Desinformation, Identitätsdiebstahl, Rufschädigung oder Manipulation verwendet werden.

So haben im Frühjahr 2023 zwei Deepfakes für ein besonderes Medienecho gesorgt: Papst Franziskus im weißen Daunenmantel (lustig!) und Donald Trump bei seiner – natürlich fingierten – Verhaftung (irritierend!). Seitdem ist vielen Menschen bewusst, dass bei im Netz verbreiteten Fotos Skepsis angebracht ist. Gleichzeitig steigt die Verunsicherung: Welchen Informationen, welchen Bildern kann man eigentlich noch trauen? Hier könnte zumindest in Deutschland die Stunde der öffentlich-rechtlichen Medien schlagen, stehen sie doch für vertrauenswürdigen Journalismus.





8.2Stehen wir vor einer neuen Spaltung der Gesellschaft?

Die Nutzung von Sprachmodellen wie ChatGPT wird künftig eine zentrale Kulturtechnik darstellen. Chatbots, die unsere Sprache beherrschen und dialogisch funktionieren, werden in den Rang von Assistenten des Menschen in seinem Alltag, im Beruf, bei seinen Entscheidungen gelangen. Wer daran nicht teilnimmt, wer Schülern zum Beispiel die Verwendung von ChatGPT in der Schule verbietet, wird dafür verantwortlich sein, dass eine Zweiklassengesellschaft entsteht: Auf der einen Seite stehen diejenigen, die es gelernt haben, KI und Chatbots

zielgerichtet als Werkzeuge einzusetzen. Sie schreiben die besseren Texte, können schneller programmieren, sind umfassender informiert, arbeiten schneller und lernen bereits früh, was Effizienz bedeutet und dass damit auch Lebensqualität verbunden ist. Denn wer möchte ernsthaft Tabellen aufwendig selbst erstellen, wenn er die Möglichkeit hat, dies die KI übernehmen zu lassen? Auf der anderen Seite bleiben jene zurück, die den richtigen, nämlich kritischen Umgang mit ChatGPT und & Co nie gelernt haben. Diese sind in zwei Gruppen einzuteilen:

- Die **Gruppe der Ahnungslosen** bildet sich aus Menschen, die wenig Schulbildung haben, sich mit prekären Jobs über Wasser halten, eingeschränkte Sprachkenntnisse besitzen oder am Rand der Gesellschaft leben. Sie bräuchten einen Mentor, der ihnen den Zugang zu Sprachmodellen verschafft und sie in dessen Bedienung und Gebrauch einführt. Dann hätten sie unter Umständen eine Chance, an ihren Lebensumständen etwas Grundlegendes zu verändern. Ohne Begleitung von außen wird dies nicht funktionieren.
- Die **Gruppe der Skeptiker** hingegen verwendet aus Prinzip kein digitales Sprachmodell, um sich damit die Arbeit zu erleichtern oder schnell eine Information zu recherchieren. Chatbots sehen sie als Teufelswerk, dessen Gefahren die restliche Menschheit bald erliegen wird. Die Technologie wird komplett und absolut abgelehnt, weil sie angeblich keinen Mehrwert bietet und die Menschen abhängig, dumm und unselbstständig macht.

Diesen Vorurteilen zu begegnen, sie zu entkräften, ist schon allein deshalb nicht ganz einfach, weil sie natürlich auch einen wahren Kern enthalten. Dieser soll im nächsten Aspekt Abschnitt besprochen werden.

8.3Macht ChatGPT uns alle dumm?

Verlernen wir jetzt endgültig selbstständiges und kritisches Denken? Überlassen wir jede Form aufwendiger und kritischer Arbeit der Maschine, dem Chatbot, der sie für uns erledigt? Brauchen wir künftig nicht mehr selbst zu programmieren, Referate selbst zu strukturieren und zu formulieren? Haben sich die Anstrengung und Bemühung um wissenschaftliche Arbeiten, das persönliche Verfassen von Mails und der Kundenkontakt nun endgültig erledigt? Kann ich alle Rezeptbücher und Lebensratgeber wegwerfen? Und wenn wir nichts mehr selber machen: Enden wir als einfältige Menschheit, die sich von Fake News manipulieren lassen und auf jeden Rattenfänger hereinfallen?

Dystopische Einschätzungen wie diese finden sich in einer Vielzahl

von Kommentaren, Interviews und sicherlich gut gemeinten Artikeln. Sie gleichen sich durchgehend in ihrer negativen und ängstlichen Sichtweise auf Chatbots und verbreiten Verunsicherung vor allem unter jenen, die grundsätzlich gegenüber technologischen Innovationen skeptisch sind (jetzt aber auch ein Smartphone nutzen) oder gern nachplappern, was ihnen in den Medien suggeriert wird.

Die von ChatGPT ausgehende Disruption in den meisten Arbeitsbereichen und vor allem in den Bildungsinstitutionen ist erkenn- und absehbar. Wir befinden uns wahrscheinlich bereits mitten darin. Die Veränderungen werden enorm sein, und die Arbeits- und Bildungswelt in 25 Jahren wird mit der heutigen nicht mehr viel gemein haben. Warum? Weil wir uns nicht gegen die Technologie stemmen können, wie das auch bisher nicht funktioniert hat. Der Buchdruck, die Elektrizität, das Autos oder später die Atombombe! Plötzlich war etwas da und es verschwand nicht mehr. Was erfunden wird, wird verwendet, egal ob der Benefit oder die Risiken überwiegen. So sind auch die künstliche Intelligenz und ihre sprachliche Variante ChatGPT zu betrachten. Wir werden diese Anwendung benutzen, einfach weil es sie gibt. Aber dümmer und unselbstständiger wird die Menschheit im Ganzen dadurch sicher nicht. Oder: Sie muss es nicht werden. Verhindern können dies unsere Bildungsinstitutionen, die ihre Schüler vorausschauend vorbereiten und ihnen beibringen, wie sie mit Chatbots als Werkzeug umgehen und sich damit selbst weiterentwickeln können.

Kommen die Schulen und Ausbildungsstätten dieser Verantwortung allerdings nicht nach, ignorieren oder verbieten sie die Verwendung von ChatGPT, dann allerdings kann es zu Verdummung in größerem Ausmaß kommen. Denn ChatGPT unkritisch und ohne Gegen- bzw. Faktencheck zu verwenden, erscheint auf die Dauer problematisch bis gefährlich.

8.4Leben mit allerlei digitalen Assistenten

Die Implementierung des digitalen Assistenten im Handgelenk oder im Unterarm? Gleich bei der Einschulung oder doch erst im Alter von zehn Jahren? Nein, das wäre ein Thema für den nächsten Science-Fiction-Roman und gehört sicher nicht hierher. Aber es streift eine Vision, die bereits in vorausgehenden Kapiteln angesprochen wurde. Chatbots geben künftig jedem Individuum die Möglichkeit, einen mitdenkenden Begleiter an seiner Seite zu haben. Denn er wird sich an die Vorlieben und Gewohnheiten des Benutzers anpassen bzw. diese erlernen und so zu einem nützlichen Alltags- und Arbeitswerkzeug werden.

Was uns momentan in unserem Alltag oft überfordert, nichts vergessen

zu dürfen, die beruflichen Anforderungen neben privaten Terminen zu stemmen, Geburtstage, Hochzeitstage, das Firmenjubiläum, das mit einem Jahr Vorlauf vorbereitet werden muss, die Abgabe der Steuererklärung, für die man mindestens vier Wochen einplanen muss, all das kann man künftig mit einem virtuellen Assistenten viel einfacher managen.

Sogar die Durchführung von Internetrecherchen kann man sich abnehmen lassen. Sucht man ein Paar neue Laufschuhe, dann gibt man dem Assistenten die Rahmendaten. Dieser wird dann passende Produkte suchen, die Preise vergleichen, und sogar, wenn gewünscht, die Bestellung selbst aufgeben. Auch Mails muss man nicht mehr selbst beantworten (man kann es aber natürlich noch). Der Assistent kann den gesprochenen Text per Spracherkennung in geschriebenen Text umwandeln und an die richtige Person verschicken. Umgekehrt muss niemand mehr Mails lesen, sondern kann sich diese vorlesen lassen. Wie praktisch!

Der Chatbot weiß auch, welche Nachrichtenkanäle von seinem User bevorzugt werden, und kann permanent die wichtigsten Nachrichten zusammenfassen und jederzeit aktuell zur Verfügung stellen. Dies ist nur eine kleine Auswahl der unzähligen Möglichkeiten, die der digitale Assistent bieten wird. Bald ist dies keine gruselige Science-Fiction mehr, sondern wird wie das Smartphone zu unserem gewohnten Alltag gehören.

8.5Klimakrise versus Energieschlucker KI?

Die KI im Allgemeinen und ChatGPT im Speziellen haben zahlreiche Auswirkungen auf unser Leben und unsere Gesellschaft, natürlich zu viel für einen kurzen Überblick, wie wir ihn hier geben wollten. Ein letzter Aspekt scheint jedoch wirklich wichtig, weil der früher oder später seinen Weg in die öffentliche Diskussion finden wird. Die KI ist nämlich ein Klimakiller par excellence! Radikale Klimaschützer fordern ein Tempolimit, der Wirtschaftsminister verordnet klimafreundliche Heizungen und wer den Flieger im Urlaub nutzt, behält das lieber für sich.

Künstliche Intelligenz und Energiesparen sind die größten Feinde überhaupt. Das Training von Sprachmodellen ist extrem rechenintensiv, benötigt überaus leistungsfähige Hardware, die Unmengen an Energie verschlingt. KI geht nicht ohne riesige Datenmengen, das Training eines neuronalen Netzwerks bedeutet, dass die Daten viele Male durch das Netzwerk laufen, um den Trainingserfolg zu steigern. Der Bedarf an

Energie dafür ist immens. Energie im großen Umfang wird auch für die Kühlung und Datenspeicherung in den KI-Trainingszentren verbraucht.

Der Gebrauch von Chat GPT durch den User, der Prompts eingibt und Output generiert, erfordert zwar nur einen Bruchteil der Energie, die beim Trainieren eingesetzt wird, hier macht sich aber der Faktor Zeit und die Anzahl der Nutzer bemerkbar. Wie passt dies zu unserer gesellschaftlichen Herausforderung Nummer eins, den Klimawandel zu bekämpfen, indem wir weniger Energie verbrauchen? Vordergründig erst einmal gar nicht. Solange fossile Energien für die Erzeugung von Strom verbraucht werden, bleibt die Rechnung negativ.

Dennoch kann KI und ChatGPT individuell und in unterschiedlichen gesellschaftlichen Bezügen dafür genutzt werden, den Energieverbrauch zu optimieren. Für jeden Einzelnen setzt dies an bei detaillierten Informationen für den eigenen Haushalt und umfasst Kochen und Waschen, den sinnvollen Gebrauch von Elektronik, den Bereich Isolierung, Beleuchtung und Heizung. Im Bereich der Forschung kann das Sprachmodell dabei hilfreich sein, Literatur schnell zu durchsuchen, Hypothesen zu formulieren oder Datenanalysen durchzuführen. Es ist auch möglich, dass man Maschinen klimabezogene Daten analysieren und Vorhersagen über zukünftige Klimaveränderungen treffen lässt. Politikern und anderen Entscheidungsträgern kann dies dabei helfen, Entscheidungen auf breiterer Informationsbasis zu treffen.

Der Energieverbrauch wird für viele Firmen ein immer größerer Kostenfaktor. Auch hier kann KI dazu beitragen, Prozesse so umzugestalten, dass damit ein geringerer Ressourcenverbrauch und weniger CO_2-Emissionen verbunden sind. Wie an anderer Stelle bereits erwähnt, wird die Landwirtschaft der Zukunft, in Zeiten von Dürren und gewaltigen Niederschlägen, auf die KI zurückgreifen müssen, um Aufwand und Ertrag in ein günstigeres Verhältnis zu bringen und gleichzeitig klimatechnisch möglichst neutral zu werden.

Ein Beitrag zum Thema KI und Klimaschutz kommt von unerwarteter Seite: Christoph Meinel, einer der führenden Informatiker Deutschlands, war langjähriger Geschäftsführer des Hasso-Plattner-Instituts und eine kritische Stimme in Sachen Klimaschutz. Er warnte vor überzogenen Erwartungen an die KI und sah neben dem Datenschutz den Energieverbrauch als wichtigsten Hinderungsgrund. Ein flächendeckender Durchbruch der KI in der Arbeitswelt scheitere daran, dass die Systeme gewaltige Rechnerkapazitäten und eine riesige Menge an Energie benötigten. Daher sei eine breite Einführung von KI für das Klima und die Erreichung der Klimaziele fatal. Seiner Ansicht nach müssten wir erst mal deutlich energieeffizientere KI-Systeme entwickeln, um den nächsten Schritt tun zu können.

8.6Virtuelle Models und Influencer: Perfekte Schönheit ohne Zicken

Wer bisher die heilen Instagram- und TikTok-Welten der Models und Influencer als wenig authentische Kopie der Realität mit dem Hang zur Idealisierung und perfekten Scheinwelt angesehen hat, kennt einen ganz neuen Trend der bildgenerierenden KI vielleicht noch nicht: Neue Influencer schießen wie Pilze aus dem Boden und erzeugen schnell Zehntausende Follower und eine immense Reichweite, ohne dass es sie in der realen Welt je gegeben hätte.

Ein paar Fotos, eine Handvoll kluger Prompts und die richtigen KI-Tools lassen virtuelle Charaktere entstehen, die werberelevante Reichweiten zu einem Bruchteil der üblichen Kosten erschaffen. Perfekte Schönheiten vor einzigartigen Traumkulissen in magischen Momenten an den schönsten Orten der Welt – statt endloser Flugstunden braucht es nur wenige Minuten in einem Hochleistungs-Rechenzentrum, um mithilfe einschlägiger KI-Tools die perfekten Illusionen am Fließband zu schaffen. Und damit werberelevante Reichweite zu einem Bruchteil der bisherigen Kosten zu generieren. Virtuelle Models und Influencer arbeiten rund um die Uhr, ohne müde zu werden, können einfach alles und machen dienstbeflissen schlichtweg all das, was sich ein zahlungskräftiger Auftraggeber wünscht. Ohne Zicken!

»Kyra« (*instagram.com/kyraonig/*) war eine der ersten virtuellen Influencerinnen auf Instagram, wird von FutrStudios betrieben und kreiert offenbar reichlich Werbeeinnahmen. »Shudu« (*instagram.com/shudu.gram/*) gilt als erstes digitales Supermodel und wird von einer virtuellen Modellagentur vermarktet, bei der man mit dem nötigen Kleingeld auch *Shootings* beauftragen kann. Die Accounts von »Kyra« und »Shudu« geben offen zu, dass es sich um virtuelle Figuren und nicht um echte Personen handelt. Wie viele Accounts virtueller Modelle nicht in derselben Weise gekennzeichnet sind und eine reale Existenz der Personen nur vorgaukeln, ist natürlich nicht bekannt. Die geschätzte Dunkelziffer ist aber hoch. Ganz anders liegt der wenig bekannte Fall der jungen Adriana Chechik, einer Pornodarstellerin, die nach einem tragischen Unfall eine Querschnittslähmung davontrug und nun ihrer bisherigen Tätigkeit nicht mehr nachgehen kann. Sie hat sich selbst ein zweites Leben als KI-Roboter geschenkt und arbeitet nun offenbar wieder in derselben Branche.

Virtuelle Models sind auf jeden Fall die angenehmere Seite der Deepfakes und nicht so gefährlich wie politische Kampagnen oder eine ganz neue Art der Deepfakes, von den die Zeitungen in Spanien unlängst berichtet haben. Dort zirkulierten gefälschte Nacktbilder von realen

Mädchen in einer Schule, die von Mitschülern offenbar ganz gezielt als Mobbing-Instrumente eingesetzt werden. Der Fall landete schnell im spanischen Justizministerium und wurde dort als *bildbasierte Gewalt* eingestuft. Bestimmt gibt es Ähnliches bereits in unserem Land, und bestimmt sind die Strafverfolgungsbehörden mit solchen neuen Delikten erst mal überfordert.

8.7 Werden Foren, Blogs und das informationsorientierte Web überflüssig?

Ein weiteres Szenario ist denkbar und sogar fast wahrscheinlich, das langfristig massive Auswirkungen auf erklärende Websites und Ratgeber-Blogs haben könnte. Diese werden nämlich zu einem großen Teil von Personen oder Unternehmen betrieben, die ein finanzielles Interesse verfolgen: Der Ratgeber für die beste Ernährung beim Fitnesstraining ist mit einem Shop für Nahrungsergänzungsprodukte verknüpft, der Blogartikel zur Reparatur einer alten Vespa führt direkt ins Werkzeug- und Teilelager eines Spezialanbieters, und der Reisetipp mit den schönsten Stränden auf Korfu enthält die Links eines Reiseveranstalters. All diese Texte und Inhalte werden künftig nicht nur von Menschen gelesen, sondern auch von der KI gefunden und verarbeitet. Irgendwann kommen dann dieselben Tipps zur Ernährung, die besagte Reparaturanleitung und die schönsten Strände Korfus als Antworten von ChatGPT oder meinem persönlichen KI-Assistenten. Ich muss nicht mehr bei Google suchen.

In der Folge verringert sich notgedrungen der messbare wirtschaftliche Erfolg der Shopanbieter und Portalbetreiber, die Investitionen in immer neuen Content lohnen sich immer weniger und irgendwann verlieren auch private Anbieter die Lust am Schreiben, wenn es außer den gefräßigen KI-Bots keiner mehr lesen will. Irgendwann wird sich der *Umweg Internet* für die Suchenden nicht mehr lohnen, und dann verschwindet langfristig ein Großteil informationsorientierter Websites. Mit bisher ungeahnten Folgen für unsere digitale Informationskultur, die wir in den letzten 20 Jahren aufgebaut und schätzen gelernt haben.

Und noch eine Gefahr droht: Schon jetzt steht zu befürchten, dass viele dieser informationsorientierten Websites immer stärker auf KI-generierte Texte umsteigen werden, die wiederum das Futter von künftigen Versionen der großen Sprachmodelle (LLMs) sein werden. Die Folge: Die heutige Vielfalt der Information wird auf einen kleinen originären Rest zusammenschrumpfen mit einem wachsenden Anteil an künstlich *wiedergekäuten* Inhalten. KI liest KI liest KI liest – eine fatale rekursive

Schleife ist die Folge. Woher nehmen neue KIs dann noch gute Informationen? Wenn kreative geistige Leistungen mehrfach den Kreislauf *Computer liest ein – Computer spuckt aus* durchlaufen haben, wird die reine Information als Destillat der Sätze dann immer besser und hochwertiger? Wohl kaum. Es ist eher zu befürchten, dass mit jeder Schleife eine Art *intellektueller Reibungsverlust* entsteht und der Informationsgehalt in den Texten zurückgeht.

Haben Sie als Kind auch *Stille Post* gespielt und den vielfachen *akustischen Reibungsverlust* erlebt? Dann haben Sie eine Vorstellung davon, was im schlechtesten Fall von einer klaren Aussage oder Information noch übrig bleibt, wenn sich die Reibungsverluste aufsummieren. Aber vor der Entwicklung von sinnvollen Lösungsansätzen für dieses Problem müssen wir erst mal ein Problembewusstsein dafür aufbauen, was sicher noch viele Jahre dauern wird.

Können wir aus all dem ein sinnvolles Fazit ziehen? Eine Prognose sei erlaubt: In 50 Jahren werden die Menschen wohl mit einem Lächeln auf unsere umwälzenden 20er-Jahre zurückblicken. Die technologische Disruption, die uns heute befremdet, beängstigt oder auch euphorisiert und antreibt, wird bis dahin bewältigt und bereits Geschichte sein. Wer sich jetzt nicht an alten Strukturen festklammert, kann sich der unausweichlichen Veränderung sicherlich besser stellen.

9Urheberrecht und Datenschutz

In diesem Kapitel wollen wir uns den Themen Urheberrecht und Datenschutz widmen und damit zwei sehr kritische Bereiche für ChatGPT ansprechen.

Beginnen wir mit dem Thema Urheberrecht und einer kleinen Einführung in die Welt des Schutzes geistigen Eigentums. Das deutsche Urheberrecht schützt nämlich das geistige Eigentum von Künstlern, Autoren und anderen Kreativen in sorgfältig ausformulierten Grenzen. Die offiziellen Regeln sind im Urheberrechtsgesetz (UrhG) festgelegt und haben eine lange Geschichte, die bis ins 19. Jahrhundert zurückreicht. Das Urheberrecht in Deutschland unterscheidet sich in einigen Aspekten von den Gesetzen in anderen Ländern, insbesondere von denen im angloamerikanischen Raum. Dazu gleich mehr.

9.1Grundlagen des deutschen Urheberrechts

Das deutsche Urheberrechtsgesetz gilt für *Werke*, also individuelle Schöpfungen auf den Gebieten der Literatur, Wissenschaft und Kunst. Dabei ist nicht jede Schöpfung automatisch geschützt. Ein Werk muss bestimmte Voraussetzungen erfüllen, um unter den Schutz des Urheberrechtsgesetzes (UrhG) zu fallen. Zunächst muss es einem der folgenden *Werkarten* angehören: literarische Werke wie Romane und Gedichte oder Artikel, musikalische Werke aller Art, Werke der bildenden Künste, inklusive Werke der Baukunst und der angewandten Kunst, Lichtbildwerke und Filmwerke sowie Darstellungen wissenschaftlicher oder technischer Art wie Zeichnungen, Pläne, Karten, Skizzen, Tabellen und plastische Darstellungen. Selbst pantomimische Werke der Tanzkunst sind urheberrechtlich geschützt.

Damit ein Werk urheberrechtlich geschützt ist, muss es als zweite Voraussetzung eine gewisse *Schöpfungshöhe* aufweisen. Das bedeutet, es muss eine individuelle Prägung vorliegen, darf also nicht lediglich eine Nachahmung bereits bestehender Werke sein. Drittens muss eine bestimmte *Gestaltungshöhe* vorliegen, die sich vom Alltäglichen, Banalen oder von bloßen Handwerksleistungen abhebt. Wichtiges Kriterium ist das Vorliegen einer kreativen Eigenleistung des Schöpfers. Damit sind Werke wie Datenbanken oder Telefonbücher von vornherein ausgeschlossen, weil sie die Anforderungen an die Schöpfungshöhe nicht erfüllen.

Das Urheberrecht ist ein Persönlichkeitsrecht, wenn es das Ergebnis einer persönlichen geistigen Schöpfung ist. Das Werk als zentraler Begriff wird als Resultat einer geistigen Tätigkeit des Urhebers verstanden und muss seine ganz individuellen Züge tragen. Nicht geschützt sind Ideen und Konzepte, Verfahren und Methoden, nur die konkrete Umsetzung oder Darstellung dieser Ideen kann, sofern sie die erforderliche Schöpfungshöhe erreicht, urheberrechtlichen Schutz genießen.

Der Urheber hat das Recht, als solcher genannt zu werden und gegen Verfälschungen seines Werkes vorzugehen. Werke sind im Übrigen automatisch geschützt, sobald diese geschaffen wurden. Eine Registrierung, wie sie in einigen anderen Ländern notwendig ist, ist in Deutschland nicht erforderlich. Das UrhG gewährt dem Urheber verschiedene Rechte, er hat das alleinige Recht, über die Nutzung seines Werkes zu entscheiden. Die Nutzung beinhaltet das Vervielfältigungsrecht, Verbreitungsrecht, Ausstellungsrecht und das Recht auf öffentliche Wiedergabe. In Bezug auf seine Verwertungsrechte kann der Urheber die Nutzungsrechte verkaufen oder lizenzieren. Zuletzt nennt das Gesetz auch moralische Rechte, die die persönliche Beziehung des Urhebers zu seinem Werk schützen, nicht übertragbar sind und immer beim Urheber verbleiben. Ohne seine Zustimmung dürfen aber private Kopien angefertigt werden, klar ausgewiesene Zitate sind ebenso erlaubt wie die Nutzung der Werke in Schulen und Universitäten zu Unterrichtszwecken.

Das Urheberrecht in Deutschland gilt lebenslang und erlischt erst 70 Jahre nach dem Tod des Urhebers. Danach wird das Werk gemeinfrei und darf von jedermann ohne Erlaubnis oder Vergütung genutzt werden.

Die Verletzung des Urheberrechts kann zu Abmahnungen und Schadensersatzforderungen führen. Besonders im digitalen Zeitalter mehren sich die Fälle von Urheberrechtsverletzungen, z. B. durch illegale Downloads oder die Verwendung von Bildern ohne Lizenz.

9.2Verwertungsgesellschaften vertreten das Urheberrecht

Die sogenannten *Verwertungsgesellschaften* bernehmen in Deutschland die Aufgabe, die Rechte von Urhebern und ausübenden Künstlern zu verwalten und durchzusetzen. Sie treten als Mittler zwischen den Rechteinhabern und den Nutzern von urheberrechtlich geschützten Werken auf und sorgen dafür, dass die Urheber angemessen vergütet werden, wenn ihre Werke genutzt werden. Verwertungsgesellschaften lizenzieren die Nutzung von Werken, sammeln dafür Gebühren ein und

verteilen diese anschließend an die Rechteinhaber. Dies betrifft z. B. das Abspielen von Musik in Radiosendern und auf öffentlichen Veranstaltungen, die Vervielfältigung von Texten für Schulbücher oder das öffentliche Vorführen von Filmen.

Die **GEMA** (Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte) ist die wohl bekannteste Verwertungsgesellschaft in Deutschland, sie vertritt die Rechte von Musikschaffenden wie Komponisten, Textdichtern und Musikverlegern. Die **VG Wort** (Verwertungsgesellschaft Wort) vertritt die Interessen von Autoren und Verlegern von literarischen Werken, von wissenschaftlichen Publikationen bis zu Zeitungsartikeln. Weitere Verwertungsgesellschaften sind die **VG Bild-Kunst** (für Fotografen und Filmemacher) und die **GVL** (Gesellschaft zur Verwertung von Leistungsschutzrechten), die sich um die Interessen von ausübenden Künstlern sowie Musik- bzw. Filmproduzenten kümmert. Die Tätigkeit dieser Verwertungsgesellschaften in Deutschland ist im Urheberrechtswahrnehmungsgesetz (UrhWahrnG) geregelt. Sie sorgen unter anderem dafür, dass Urheber für die Nutzung ihrer Werke in verschiedenen Kontexten eine angemessene Vergütung erhalten, und haben eigene Tarifstrukturen dafür aufgebaut. Sie erheben Pauschalabgaben auf Geräte, die zur Vervielfältigung von Werken genutzt werden können, wie Drucker, Kopierer und Speichermedien.

Die meisten Länder schützen geistiges Eigentum, die konkrete Gestaltung der Schutzrechte variiert aber von Land zu Land. Im englischen (UK) bzw. amerikanischen (USA) Urheberrecht gibt es im Detail einige bemerkenswerte Unterschiede: So können in den USA Werke von Unternehmen oder anonym veröffentlichte bis zu 120 Jahre ab Erstellung geschützt sein. Die größten Unterschiede gibt es bei den formalen Anforderungen. Wir haben schon gehört, dass es in Deutschland keine Registrierungspflicht gibt und der Schutz automatisch mit der Schaffung des Werks entsteht. Dasselbe gilt in UK, in den USA kann eine Registrierung beim *U.S. Copyright Office* notwendig sein, um dort Schadensersatzklagen wegen Urheberrechtsverletzung einreichen zu können. In den USA gibt es auch das Konzept des *Fair Use*, das bestimmte Nutzungen von urheberrechtlich geschützten Werken ohne Zustimmung des Urhebers erlaubt, entscheidend ist der Kontext und die Natur der Nutzung.

Das Urheberrecht und die starken Verwertungsgesellschaften empfinden viele als sehr rigoros und als eine Beschränkung der eigenen kreativen Entfaltungsmöglichkeiten. Aber was würde passieren, wenn es kein Urheberrecht gäbe? Dies hätte tatsächlich weitreichende Folgen für Künstler, Produzenten, Konsumenten und die gesamte Kreativ- und Kulturindustrie, unsere kulturelle Vielfalt würde merklich leiden. Für

Kreative gäbe es mangelnde Anreize, denn wenn Künstler, Schriftsteller und andere Kunstschaffende nicht in der Lage wären, von ihrer Arbeit zu profitieren, könnte dies ihre Motivation erheblich verringern. Ein Mangel an wirtschaftlichen Anreizen würde zu einem Rückgang der Qualität kreativer Werke führen, da weniger Mittel für Produktion, Forschung und Entwicklung zur Verfügung stehen. Verlage, Musiklabels, Filmproduzenten und andere Unternehmen in der Kreativbranche würden sicherlich erhebliche Verluste erleiden, da sie nicht mehr exklusiv von den Werken profitieren könnten, in die sie investiert haben. Wie schon gesagt, droht ein ernster Verlust kultureller Vielfalt, unter der wir alle zu leiden hätten. Ein starkes Urheberrecht, effektive Verwertungsgesellschaften mit gesetzlichem Mandat und wirksame Rechtsmittel zur Durchsetzung der Schutzrechte sind in unser aller Interesse.

9.3Der größte Diebstahl des Jahrhunderts?

Zurück zu ChatGPT: Der Transformer braucht große Mengen an Trainingsdaten, also sehr, sehr viele Texte aus allen Bereichen, von der Literatur bis hin zu Gebrauchstexten aller Art. Ein großer Anteil des Trainingspools ist der *Common Crawl*, also Texte, die im Internet gefunden und in einer großen Datenbasis gesammelt wurden. Auch von *Books* ist die Rede, ohne dass wir erfahren hätten, ob oder wie lange die Autoren dieser Bücher bereits tot sind. Wir müssen also davon ausgehen, dass ein großer Teil dieser Texte von Privatpersonen geschrieben wurde, für die selbst in den USA automatisch das Urheberrecht oder eine Art Copyright gilt. ChatGPT und alle anderen GPTs (Generative Pretrained Transformer) brauchen diese Texte, um selbst neue Texte erzeugen zu können. Wird damit das Recht der Autoren am eigenen Werk im ganz großen Stil missachtet, das Urheberrecht quasi mit Füßen getreten?

Das ist eine Frage der Sichtweise, von denen es immer wenigstens zwei gibt: Weil das Sprachmodell Texte nicht direkt übernimmt und wiedergibt, ist eine Verletzung von bestehenden Schutzrechten nicht gegeben und auch nicht nachweisbar. Und weil so viele Texte in das KI-Modell eingeflossen sind, *verdünnt* sich eine unterstellte unerlaubte Nutzung von geistigem Eigentum anderer bis tief in den Nanobereich hinein. Also bis weit unterhalb jeder Bagatellgrenze mit entsprechend niedrigem Schaden und nicht nennenswertem Streitwert im Falle einer Klage. Und schließlich haben alle Autoren auch andere Autoren gelesen, bevor sie eigene Texte erstellen konnten. Also doch kein Diebstahl, sondern ein ganz üblicher Kreislauf von *geistigem Recycling*, das wir über

Generationen bereits betreiben und das vielleicht gar eine Voraussetzung für die Entwicklung der Menschheit darstellt?

Aber man könnte auch andersherum argumentieren, dass OpenAI und andere Hersteller auf Milliarden von *Nano-Verletzungen* ein Tool aufgebaut haben, mit dem man – zumindest irgendwann in der Zukunft – richtig großes Geld verdienen will. Wenn man es aus dieser Perspektive betrachtet, dann sind Transformer systematische Ausbeuter geistig Schaffender, verleiben sich widerrechtlich und ohne Vergütung die Texte anderer ein, um daraus eine neue Leistung zu erstellen und selbst zu vermarkten. Also der größte Diebstahl des Jahrhunderts? Gegen den alle anderen dreisten Diebstähle wie Kavaliersdelikte anmuten?

Man kann sich eine der beiden Argumentationsrichtungen – Diebstahl oder Recycling – zu eigen machen, jeder wohnt eine nachvollziehbare Logik inne. Um den möglichen Ausreden gleich mal den Wind aus den Segeln zu nehmen: Den hier unterstellten systematischen Diebstahl als Zitate zu rechtfertigen, ist nicht zulässig, dazu fehlt nämlich ihre Kennzeichnung. Auch ein wissenschaftlicher Nutzen ist nicht gegeben, weil der Chatbot in kommerziellen Kontexten auftaucht und explizit für diese gemacht wurde. Trifft vielleicht das amerikanische Konzept des *Fair Use* zu? Schwer zu sagen, allerdings spricht der Bot ja auch Deutsch, was man wiederum nur aus deutschen Texten lernen kann, die wohl zu einem großen Teil von deutschen Autoren unter dem Schutz unseres strengen Urheberrechts entstanden sind.

9.4Geistiges Recycling, Freiheit der Information und Hackerethik

Lassen wir die Diebstahl-Theorie mal einen Moment aus den Augen und prüfen, was dagegenspricht und der KI legitimen Zugriff auf alles erlaubt, was in Textform irgendwo zu finden ist. Den natürlichen Kreislauf des *geistigen Recyclings* hatten wir schon angesprochen. Kinder lernen die Sprache von den Eltern, Schüler aus Büchern und Texten von klugen Leuten, die im Bildungskontext von Gesetz wegen auf ihr Urheberrecht verzichten müssen. Junge Wissenschaftler lernen vom Doktorvater und wälzen tonnenweise Literatur, bis erste Texte veröffentlicht werden. Das ist eigentlich alles ganz normal und wird seit Jahrhunderten systematisch und im ganz großen Stil betrieben. Eigentlich ist es sogar die Voraussetzung dafür, dass neue Werke überhaupt entstehen können. Warum sollten für lernende Maschinen andere Regeln gelten, zumal die KI ja offen zugibt, menschliches Lernen nachbilden zu wollen?

Die Freiheit der Information und ihre allgemeine Zugänglichkeit wird bei Google sehr hochgehalten. Das Internet war ein Katalysator für die Idee, dass Information frei, kostenlos und stets für alle zugänglich ist, auch weil die Kosten, sie öffentlich zugänglich machen, immer weiter sinken. Die Forderung nach freien Informationen für alle ist sogar noch deutlich älter als das Silicon Valley, die Idee des möglichst uneingeschränkten Informationsaustauschs zum Wohle der Menschheit ist nichts weniger als der Grundgedanke der Wissenschaft. Und lässt sich mindestens bis Platon und Aristoteles zurückverfolgen. Hacker haben Ende der 90er-Jahre diesen Gedanken der unbeschränkten und gleichzeitig kostenlosen Information für alle zum ethischen Grundprinzip der digitalen Revolution erklärt. Dieser schöne Kerngedanke hat Perlen wie Wikipedia, Open-Source-Software wie Linux und öffentliche Code-Bibliotheken hervorgebracht, die heute selbstverständlich sind.

Ein paar Widersprüche bleiben: Texte sind also frei zugänglich? Freie Texte sind dann Trainingsdaten für die KI? So weit ist das auch nachvollziehbar! Sobald freie Texte aber zu Trainingsdaten geworden sind, heißen sie plötzlich *Technologie*! Und werden dann nicht mehr öffentlich zugänglich gemacht, obwohl sie ja frei sind. Christian Stöcker, Professor an der Hochschule für Angewandte Wissenschaften Hamburg und Spiegel-Kolumnist, nennt das *Informationsasymmetrie* und bemerkt, dass dies Leute reich machen kann, wenn sie auf der richtigen Seite sitzen. Was erst noch frei verfügbar war, kann doch nicht plötzlich streng geheim sein? Unternehmen wie Microsoft und Google arbeiten in wichtigen Zukunftsbereichen und haben die reichsten Männer der Welt hervorgebracht. »Und ChatGPT ist nur deshalb ein so mächtiges Werkzeug, weil es Unmengen von Information verdaut hat, die sogenannten Trainingsdaten.«[7]

9.5Wem schadet die generative KI?

Werden also Urheber, Autoren und Texter auf der ganzen Welt durch ChatGPT und andere KI-Tools geschädigt? Wenn nicht auf direktem Wege, dann indirekt auf jeden Fall, denn der Chatbot nimmt ihnen unter Umständen die Quelle ihres Broterwerbs. Autoren und Texter haben ChatGPT nicht erschaffen, ihn aber die Sprache gelehrt, und sie sind die Ersten, deren Fähigkeiten radikal an Wert verlieren. Fotografen und Bilddatenbanken geht es nicht anders, wenn Bilder künftig mit Klicks erschaffen werden. Während die Macher der Tools davon profitieren und immer weitere Milliarden hinzuverdienen, gehen die kleinen Lieferanten des Mastfutters leer aus. Diese negative Sichtweise kann man durchaus

auch einnehmen und es wird schwierig sein, diese komplett zu widerlegen. Was ist auf diesem Hintergrund eigentlich aus dem Gründungscredo von OpenAI geworden, dass KI allen Menschen zugutekommen soll?

Ist der Streik der Copywriter und Drehbuchautoren in Hollywood, die den Verzicht auf künstliche Intelligenz wenigstens in ihrem Bereich fordern, eine geeignete Antwort? Und was tun wir im eigenen Land? Das deutsche Urheberrecht ist trotz seines Detailreichtums machtlos, weil ein konkreter Diebstahl geistigen Eigentums nicht nachweisbar wäre und weil *Nano-Verletzungen* dort nicht genannt sind. Die Verwertungsgesellschaften sind plötzlich zahnlos gewordene Tiger. Sollen wir die VG Wort eine Zeilengebühr auf die Feder von ChatGPT erheben lassen? Oder eine ganz neue Gesellschaft im Stile der GEMA gründen? Die Diskussion steht erst am Anfang, aber die Sensibilität für diese Fragen ist längst vorhanden. Vielleicht müssen wir sogar unseren Begriff von Urheberschaft erweitern, um in Zukunft allem und allen gerecht zu werden.

Im Zusammenhang mit dem Urheberrecht sind wenigstens zwei weitere Fragen interessant:

- Sind die von ChatGPT generierten Texte urheberrechtlich geschützt?
- Können KI-Texte Urheberrechte Dritter verletzen und wer haftet dann?

Ein Blick in die AGB von OpenAI zeigt, dass man sich mit diesen Fragen bereits auseinandergesetzt hat: Die Nutzenden seien einmal für den Input, aus dem ChatGPT dann neue Texte generiert, selbst verantwortlich. Außerdem räumt OpenAI den Nutzenden alle Rechte ein, die an Texten entstehen, die durch die KI generiert wurden. Für die Compliance des Contents seien die Nutzenden natürlich selbst verantwortlich.

Dabei sind KI-Texte nach dem Urheberrechtsgesetz gar nicht schutzfähig, da keine persönliche geistige Schöpfung vorliegt und das Urheberrecht nur für menschliches Schaffen gilt. Damit entsteht bei der Erzeugung von KI-Texten gar kein Urheberrecht, das von OpenAI an den Nutzer lizenziert werden könnte. Und was sagt der Chatbot (GPT-4) selbst dazu?

```
>>Verstoßen deine Texte gegen das Urheberrecht?
Nein, die Texte, die ich generiere, verstoßen nicht gegen das
Urheberrecht. Sie sind originale Kreationen, die in Echtzeit auf
Basis des von mir trainierten Modells und der mir gestellten
Fragen erzeugt werden. Ich kopiere oder reproduziere nicht exakt
```

```
Inhalte aus urheberrechtlich geschützten Quellen. Stattdessen
verwende ich das Wissen, das mir während des Trainings vermittelt
wurde, um Antworten zu generieren.
OpenAI hat mich auf der Grundlage einer riesigen Menge an Daten
trainiert, um eine breite Vielfalt von Anfragen zu beantworten.
Mein Ziel ist es, Informationen bereitzustellen und Fragen zu
beantworten, ohne dabei urheberrechtlich geschütztes Material zu
kopieren oder zu verletzen. Es ist jedoch wichtig zu beachten,
dass Nutzer, die auf meinen Antworten aufbauen oder sie
verwenden, sicherstellen sollten, dass sie keine Urheberrechte
oder andere rechtliche Bestimmungen verletzen, insbesondere wenn
sie den Inhalt für kommerzielle Zwecke nutzen möchten.
```

Folgt man dieser Argumentation »verwende ich das Wissen, das mir während des Trainings vermittelt wurde, um Antworten zu generieren«, ist es genau das, was jeder Autor tausendfach selbst tut und wir vorher als *geistiges Recycling* bezeichnet haben: Auch unsere menschliche Intelligenz braucht eine Trainingsphase, in der wir die Sprache lernen und uns Wissen aneignen, das dann in unsere Werke einfließt. Mit dieser durchaus geschickten Antwort hat sich der Bot eigentlich selbst von allen Vorwürfen freigesprochen.

9.6Ist ChatGPT datenschutzkonform?

Datenschützer sind wie Verwertungsgesellschaften nicht sehr beliebt, wahrscheinlich deshalb, weil sie ihre Arbeit so gut und so gründlich machen, dass wir den Schutz schon für etwas völlig Normales halten. Datenschützer sind im Allgemeinen nicht sehr gut auf den Chatbot zu sprechen, das liegt unter anderem auch daran, dass die Technologie in den USA entwickelt wurde und heute noch durch US-Firmen kontrolliert wird. Die Technologie hinter ChatGPT ist eine Blackbox, Datenschützer fordern mit größtmöglicher Transparenz aber genau das Gegenteil. Und allein die immense Menge der Trainingsdaten, die man auch deswegen nicht angeben möchte, weil kein Mensch die Liste je durchlesen könnte, erzeugt ein ziemliches Unbehagen bei Datenschützern. Und was passiert eigentlich mit den Daten, die ich während der Interaktion mit dem Chatbot in den Slot eingebe? So lange die Verarbeitung und Verwendung eingegebener Daten nicht vollständig geklärt ist, wird kein seriöser Datenschützer je seinen vollständigen Frieden mit dem KI-Tool machen können.

Schon aufgrund ihrer immensen Komplexität ist und bleibt die datenschutzrechtliche Beurteilung von KI eine echte Herausforderung.

Grundfragen des Datenschutzes und der DSGVO klingen so oder ähnlich:

- Woher stammen die Daten, mit denen das System trainiert wurde?
- Wie arbeiten die Algorithmen hinter der automatisierten Datenverarbeitung?
- Werden eingegebene Daten an Dritte mit kommerziellen Interessen weitergegeben?

Wir wissen heute, dass ChatGPT neben seinem Textkorpus auch Daten aus den Benutzereingaben für das Training nutzt. Wenn Nutzer persönliche Daten in das System eingeben, sind persönliche und sensible Informationen potenziell Teil des Systems und könnten als Information in den Antworten ausgegeben werden.

Aufgrund der gezeigten Mängel und fehlender Transparenz durch den Hersteller OpenAI ist heute und in der nächsten Zeit keine vollständige Datenschutzkonformität des Systems zu erwarten und wohl auf Dauer auch nicht zu gewährleisten. OpenAI will selbst keine Garantie für die Sicherheit der von Nutzern eingegebenen Daten und Informationen übernehmen. Unternehmen, die ChatGPT im kommerziellen Kontext nutzen und für ihre Zwecke z. B. in der Kundenkommunikation oder im Support einsetzen wollen, sollten dies berücksichtigen. Nutzer im privaten Umfeld sollten selbst für einen angemessen Datenschutz und für eine Reihe von Sicherheitsmaßnahmen sorgen. Die Eingabe persönlicher oder sensibler Informationen sollte tunlichst vermieden werden, außerdem wird die Nutzung des Inkognitofensters bzw. privaten Modus im Browser empfohlen und die Verwendung eines VPN-Clients. Durch den privaten Modus werden keine Daten zum Browserverlauf und keine Eingaben gespeichert, ChatGPT könnte auf solche Informationen also auch nicht zurückgreifen. Sichere Internetverbindungen über HTTPS-Verbindung sind ohnehin selbstverständlich.

9.7 Risiken beim Einsatz von ChatGPT

»Vorsicht: Der Einsatz von ChatGPT verstößt aktuell gegen den Datenschutz«. So titelt die IT-Recht Kanzlei (*it-recht-kanzlei.de*) in einem Beitrag vom Juli 2023. ChatGPT stehe im Kreuzfeuer der Datenschützer, es drohe sogar ein Verbot in Deutschland. Die strikten Vorgaben der Datenschutz-Grundverordnung (DSGVO) sehen das Tool mindestens aus vier Gründen als problematisch an:

Es enthalte *personenbezogene Daten*, die *beim Lernen* nicht systematisch ausgespart wurden, die KI selbst kann personenbezogene Daten nicht sicher von nicht personenbezogenen Daten unterscheiden. Die

Verarbeitung personenbezogener Daten erfordert aber die Zustimmung der Betroffenen als Rechtsgrundlage, die natürlich nicht vorliegt. Der zweite Grund seien personenbezogene Daten bei *Eingaben von Nutzern*, die natürlich ebenfalls nicht systematisch verhindert werden (können). Wer personenbezogene Daten an den Anbieter von ChatGPT weitergibt, braucht dazu ebenfalls die Zustimmung der Betroffenen. Personenbezogene Daten könnten auch zum *Weiterlernen* der KI genutzt werden, das wurde von OpenAI nicht explizit ausgeschlossen.

Wer solche Daten verarbeitet, ist zur *Transparenz* bei der Verarbeitung von personenbezogenen Daten verpflichtet und die betroffenen Personen müssten darüber informiert werden, was schon wegen der Menge und Komplexität der Verarbeitungsvorgänge einer KI praktisch unmöglich ist.

Die italienische Datenschutzbehörde *Garante per la protezione dei dati personali* hat im März 2023 die Verwendung von ChatGPT in Italien aus datenschutzrechtlichen Gründen untersagt, weil ChatGPT personenbezogene Daten unrechtmäßig gesammelt habe. Außerdem gebe es keine gesicherte Altersverifikation, um Kinder vom Gebrauch des Systems abzuhalten. OpenAI wurde aufgefordert, kurzfristig Maßnahmen zu ergreifen, andernfalls drohe eine Strafe in Millionenhöhe. OpenAI hatte sich beeilt, mehrere der geforderten Änderungen vorzunehmen, um die Einhaltung der europäischen Datenschutzbestimmungen zu gewährleisten, inklusiv einer Altersprüfung für inländische Nutzer. Ende April kam die Entwarnung und seitdem darf ChatGPT in Italien wieder genutzt werden. Auch deutsche und französische Datenschützer haben Bedenken geäußert, die EU will im Rahmen ihrer KI-Verordnung alle Anbieter dazu verpflichten, urheberrechtlich geschütztes Material, das zum Training von KI-Systemen verwendet wird, offenzulegen. KI-Tools sollen dann mit einem Risikoniveau zwischen *minimal* bis *inakzeptabel* klassifiziert werden.

Ein weiteres und generelles Problem besteht aus Sicht des europäischen Datenschutzes immer dann, wenn Daten in *unsicheren Drittländern* verarbeitet werden. Als unsicher gelten alle Staaten, die kein mit der EU vergleichbares Datenschutzniveau gewährleisten. Also auch die USA, wo wahrscheinlich sehr viele Server stehen, die OpenAI zum Betrieb der Software nutzt. In diesem Zusammenhang fordert das Datenschutzrecht auch den Abschluss eines *Auftragsverarbeitungsvertrags* nach anerkannten EU-Standards.

Die IT-Recht Kanzlei rät ihren Mandanten daher zu größter Vorsicht, jeder Einsatz der Software berge schwer abschätzbare Risiken. Gerade wenn Händler und Unternehmen ChatGPT als Chatbot in der Kundenkommunikation oder in anderen Anwendungen einsetzen, sollten diese unbedingt die aktuellen datenschutzrechtlichen Entwicklungen beobachten.

10Sicherheit, Kontrolle und Missbrauch

Im vorigen Kapitel wurde erstmals das Thema Sicherheit angesprochen, wobei es noch vorwiegend um die Sicherheit desjenigen ging, der ChatGPT selbst einsetzt. Gibt es auch Gefahren und Sicherheitsrisiken, wenn andere den Chatbot verwenden? Wie sieht es mit der Kontrolle aus, wenn die Fragen den gesetzlichen Horizont der Datenschützer sogar noch überschreiten? Und welche Missbrauchsmöglichkeiten ergeben sich aus dem Potenzial generativer KI für uns alle als Gesellschaft? Wer könnte am Ende die Macht der KI sinnvoll beschränken und in einer Weise bändigen, die den Missbrauch wirkungsvoll verhindert, den positiven Entwicklungen aber nicht entgegensteht?

10.1Ein Zauberlehrling, der die Geister nicht mehr loswird

In seiner Ballade *Der Zauberlehrling* erzählt uns Johann Wolfgang von Goethe die Geschichte eines jungen Lehrlings, der in Abwesenheit seines Meisters beschließt, dessen magische Kräfte zu probieren. Er spricht einen Zauber aus und erweckt damit einen Besen zum Leben, damit dieser für ihn Wasser in eine Wanne trägt. Anfangs ist der Lehrling begeistert von seiner eigenen Fähigkeit, die lästige Arbeit auf magische Weise erledigen zu lassen. Doch als die Wanne voll ist und der Besen nicht aufhört, Wasser zu holen, bemerkt der Lehrling, dass er den Gegenzauber nicht kennt und den Besen nicht mehr stoppen kann. In seiner Verzweiflung versucht der Lehrling, den Besen mit einer Axt zu zerstören, doch aus jedem Stück entsteht ein neuer Besen, und die Flut geht unaufhörlich weiter. Erst als der Meister zurückkehrt, wird der Zauber beendet und die Ordnung wiederhergestellt. Der Lehrling wird zurechtgewiesen und belehrt, dass man keine Kräfte rufen darf, die man nicht sicher zu kontrollieren weiß.

Die Relevanz der Ballade für die moderne Informatik ist ziemlich naheliegend: Die Geschichte des Zauberlehrlings kann als Metapher für unsere heutige Technologie und insbesondere für künstliche Intelligenz gesehen werden. Wie der Zauberlehrling, der mit Kräften spielt, die er nicht vollständig versteht und kontrollieren kann, so entwickeln Menschen fortschrittliche KI-Systeme, die in manchen Fällen unvorhergesehene oder unkontrollierbare Handlungen ausführen können. Handlungen, an die wir nicht gedacht haben und die keiner

gewollt hat. In der Entwicklung immer fortschrittlicherer Werkzeuge und immer stärkerer KI-Systeme dürfen wir unsere ethische Verantwortung niemals aus den Augen verlieren. Neben der Begeisterung angesichts immer neuer Potenziale brauchen wir eine respektvolle Vorsicht im Umgang mit leistungsstarken Technologien als starkes Gegengewicht. Ohne ausreichende Kontrollmechanismen oder Verständnis für die Potenziale einer Technologie kann ihre Anwendung unerwünschte oder sehr schädliche Ergebnisse haben.

Beim Training großer neuronaler Netzwerke können Modelle manchmal unerwartete Verhaltensweisen zeigen oder Voreingenommenheiten aus den Trainingsdaten übernehmen (siehe Microsoft Tay). Ohne sorgfältige Überwachung und verantwortungsvolle Kontrolle können solche Systeme potenziell schädliche Ausgaben erzeugen oder Informationen verbreiten, die nicht korrekt sind und im Widerspruch zu unseren gesellschaftlichen Werten stehen (wie die Leugnung des Holocausts). Die Ballade vom Zauberlehrling sollten wir als Warnung betrachten, Fortschritte und Innovationen nicht blind und gierig voranzutreiben, sondern uns rechtzeitig mit den potenziellen Folgen unseres Handelns auseinanderzusetzen und vor der Implementierung über robuste Sicherheitsvorkehrungen ernsthaft nachzudenken.

10.2Das Moratorium für die KI-Entwicklung

Heute trägt der Zauberlehrling den Namen *Elon Musk*, einer der Gründerväter von OpenAI, der dem Unternehmen jedoch schon bald wieder den Rücken kehrte. Nur wenige Monate nach der Veröffentlichung von ChatGPT schlagen er und namhafte Experten eine sechsmonatige Entwicklungspause von Systemen wie ChatGPT vor. Das Moratorium für KI-Entwicklung solle der öffentlichen Diskussion Zeit geben und eine breite Diskussion über eine mögliche Regulierung ermöglichen, bevor diese von immer neuen Entwicklungen schnell wieder eingeholt würden. Die Zeit des Moratoriums solle genutzt werden, um Regelwerke zu schaffen, die neue KI-Systeme noch vor ihrer Entwicklung sicher und die Risiken überschaubar machen. Notfalls solle der Zeitrahmen auch verlängert werden, die öffentliche Meinungsbildung und das uns alle schützende Recht hätten schließlich Vorrang.

Die in einem offenen Brief der gemeinnützigen Organisation *Future of Life* erhobenen Forderungen wurden von namhaften Persönlichkeiten unterzeichnet, mehr als 30.000 Unterstützer haben sich mittlerweile

online eingetragen[8]. Generative KI wie ChatGPT berge eine ganze Reihe von Gefahren, da sie menschliche Interaktionen perfekt simulieren könne. Anhand weniger Stichworte entstünden massenhaft glaubwürdige Texte oder täuschend echte Bilder, die von der Wirklichkeit völlig entkoppelt seien. Selbst die Entwickler dieser Programme würden deren Äußerungen teilweise schon nicht mehr verstehen und auch nicht wirksam kontrollieren können. Man könne damit öffentlich zugängliche Informationskanäle mit jeglicher Propaganda und sämtlichen Unwahrheiten fluten und durch Desinformation politische Entscheidungen beeinflussen. Desinformation und Propaganda gehörten zu den größten Gefahren, die von einem eloquenten und unermüdlichen Chatbot ausgehen. Dazu gleich noch mehr.

10.3Wenn KI-Systeme aus dem Ruder laufen

Sind existierende KI-Systeme eigentlich schon aus dem Ruder gelaufen? Jenseits dessen, was sich als Fiktion im Kino oder als Roman gut verkaufen lässt? In einigen Quellen wird der »Flash Crash« von 2010 (plötzlicher Börsenabsturz) als ein solches Ereignis genannt, offenbar steckte aber doch menschliche Absicht dahinter. Pannen in Empfehlungs-Algorithmen gab es bei YouTube und Facebook schon mehrmals, das war aber nichts wirklich Gravierendes.

Weil viele KI-Anwendungen ganz explizit ohne menschliche Überwachung selbstständig Entscheidungen treffen und entsprechend robust handeln sollen, kann es durchaus passieren, dass so ein System aus reinem Selbsterhaltungstrieb außer Kontrolle gerät und absurde Dinge tut. Um das zu veranschaulichen, müssen wir hier auf ein fiktives Beispiel von Jerry Kaplan zurückgreifen:

In meiner Stanford-Vorlesung zur Ethik und den Auswirkungen der KI erzähle ich die Geschichte von Curbside Valet, einem fiktiven Unternehmen, dessen Gründer einem bedauerlichen Unfall zum Opfer fällt. Seine robotischen Gepäckträger waren sein größter Stolz, vor allem seine geschickte Konstruktion für die berührungslose Bedienung. Sein Vertrag mit dem Flughafen von San Francisco sieht vor, dass Fluggäste ihr Gepäck vor dem Abflug in einem seiner niedlichen Karren verstauen und so einchecken können. Der gut verschlossene Karren transportiert das Gepäck dann sicher zur Fluggesellschaft. Und falls die Akkus schwächeln, wird automatisch eine Steckdose in einer dunklen Ecke angesteuert – natürlich möglichst in ruhigen Zeiten –, um ihn wieder aufzuladen. Im Störungsfall wird eine beauftragte Werkstatt informiert, die sich um die Reparatur vor Ort kümmert. Dabei werden Diagnosedaten und der genaue Standort übertragen. Die Einkünfte

werden automatisch einem PayPal-Konto gutgeschrieben, die Zahlungen für die Werkstatt (und alle anderen Ausgaben) werden automatisch nach Erhalt der elektronischen Rechnung geleistet. Übersteigt der Gewinn einen bestimmten Schwellenwert, werden weitere Roboter bestellt und aktiviert.

Nachdem der Gründer unfallbedingt eines viel zu frühen Todes stirbt, läuft das beliebte System noch viele Jahre unbeaufsichtigt weiter, bis der Flughafen selbst der wachsenden Beliebtheit der Senkrechtstarter zum Opfer fällt, die viel weniger Platz benötigen und daher viel näher am Stadtzentrum starten und landen. Schließlich wird der Strom abgestellt, sodass die Roboterkarren immer weiter fahren müssen, um zumeist in der Nacht eine funktionierende Steckdose zu finden. Zunächst schmunzeln die Bewohner des nahe gelegenen San Bruno über die rollenden Dinger, die automatisch nach Einbruch der Dunkelheit an Steckdosen in Garagen neben Terrassen andocken, aber irgendwann wird es zu viel: Hohe Stromrechnungen, verschreckte Haustiere und zertrampelte Vorgärten führen dazu, dass ein Team von Roboterjägern angeheuert wird, um die mechanischen Missetäter einzusammeln und zu verschrotten.[9]

Diese Geschichte ist wie gesagt reine Fiktion. Offenbar können autonome Geräte, die mit künstlicher Intelligenz gesteuert werden, zu unbeabsichtigten Komplikationen führen. Horrorvisionen und Gruselgeschichten von KI-Systemen, die die Singularität erreicht haben und dann mit Geisterstimme die Weltherrschaft übernehmen, gibt es genug. Diese kleine Geschichte von Jerry Kaplan ist deutlich realistischer und impliziert auch ein paar Sicherheitsschranken, die man schon in der Entwicklung ganz einfach hätte einbauen können. Das Risiko einer sich ungewollt verselbstständigenden KI kann man mit ein paar einfachen Regeln minimieren und wir sollten diese Regeln als verbindliche Standards formulieren. Das ist auch Teil des offenen Briefes von Elon Musk und anderen.

Teilweise gibt es solche *Fallback-Lösungen* auch schon: Drohnen, die so weit fliegen, dass die Funkverbindung zwischen Fluggerät und Steuergerät abreißt, befolgen eine Routine, bei der die Drohne wieder in eine definierte Neutralstellung zurückkehrt, also zum letzten Startplatz zurückfliegt. Im Flughafen-Beispiel könnte man das Ausbleiben von neuen Befehlen oder Eingaben nach einer gewissen Zeit wie eine unterbrochene Funkverbindung bewerten und eine vorher definierte *Neutralstellung* einnehmen: Die Geräte hören auf zu arbeiten, sammeln sich an einem bestimmten Ort oder gehen einfach in eine Parkposition. Ohne Funkverbindung bzw. außerhalb der definierten Anwendungsgrenzen müssen Systeme in eine *sichere Betriebsart* zurückkehren und in dieser Warteposition verharren.

Bei intelligenten Staubsaugern oder Rasenmähern hält sich ein möglicher Schaden wohl in engen Grenzen, eine umgefallene Vase und

eine irritierte Hauskatze fallen nicht ins Gewicht. Nicht immer hilft es aber, im Zweifelsfall unser Fahrzeug, den Rasenmäher oder Roboter in die definierte *Neutralstellung* zurückzusenden, in der sie keinen Schaden anrichten können. Bei selbstfahrenden Autos kann es erheblich gefährlicher werden, hier müssen deutlich höhere Anforderungen an die Sicherheit gestellt werden. Wenn im Falle einer Notbremsung der KI nur noch die Wahl bleibt, mit einer Ausweichbewegung entweder zwei ältere Damen auf einer Parkbank oder ein Schulkind auf dem Fußweg zu überfahren, muss die KI eine Wahl treffen. Weil diese Systeme nicht über Werte und Moral verfügen, können sie solch ein Dilemma auch nicht selbstständig entscheiden. Wir müssen sie so konstruieren, dass sie sich in allen Lebenslagen auf ethisch annehmbare Weise verhalten. Was nichts weniger bedeutet, als dass wir alle diese Fälle vorhersehen und Befehle dafür hinterlegen müssen. Oder Regeln, aus denen sich diese Befehle klar ableiten lassen. Und als Gesellschaft diese Regeln festlegen – was wohl der schwerste Teil der Aufgabe ist. Menschen in der gleichen Situation entscheiden hingegen schnell und spontan, haben sicherlich nicht alle Informationen und können in der akuten Stresssituation auch nicht alle Konsequenzen vorhersehen und abwägen. KI-Systeme in autonomen Fahrzeugen brauchen gesellschaftlich akzeptierte Leitplanken.

10.4Kontrollverlust – nur gefühlt oder schon echt?

KI-Systeme können menschliche Entscheidungen begleiten und absichern, sie sollen mehr oder weniger erfahrene Fachberater bei ihrer Arbeit unterstützen. Solche Tools werden zuerst *gedacht*, dann entwickelt, mit Daten gefüttert und schließlich implementiert. Dann sind sie da und werden auch eingesetzt. Gut, wenn Mensch und Maschine in der Praxis dann zum selben Ergebnis kommen: Die Bank kann ruhigen Gewissens den Immobilienkredit genehmigen, das Unternehmen einen bestimmten Bewerber einstellen und der Arzt bekommt grünes Licht, eine seltene Therapie zu verordnen.

KI-basierte Entscheidungsunterstützungssysteme (ja, so heißen die wirklich) können dem menschlichen Empfinden aber auch widersprechen. Und was ist dann zu tun? Der freundliche Bankberater weiß es doch eigentlich besser als seine Software, der Personaler hat mehr Vertrauen zum Bewerber als in das neue Tool und der Arzt schließlich genug Erfahrung, um dem Computer zu widersprechen. Aber das Risiko, eine von der Empfehlung abweichende Entscheidung zu treffen, ist hoch, die persönliche Verantwortung und mögliche

Haftungsrisiken sind nicht zu übersehen. Wir müssen davon ausgehen, dass solchen KI-Systemen schon aus Angst vor einer eigenen Fehlentscheidung und der sich daraus ergebenden Verantwortlichkeit immer weniger oder überhaupt nicht widersprochen wird. Aus Entscheidungsunterstützungssystemen werden dann schleichend Entscheidungssysteme.

Gerade wenn sich solche KIs dann selbst immer weiterentwickeln, droht ein *Kontrollverlust*, da immer weniger Menschen noch wissen, was die KI wann und warum tut. Widerspruch wird immer riskanter, die Möglichkeiten zur eigenen Entscheidung schwinden. Beim Wort *Kontrollverlust* geht es also nicht um dystopische Horrorszenarien, sondern um Alltagserfahrungen, die zunehmen werden: Der Kontrollverlust ist dann sehr real, wenn der Mitarbeiter einer Bank mit den Schulten zuckt, der freundliche Personaler sein persönliches Bedauern ausdrückt und der Arzt eine möglicherweise lebensrettende Behandlung verweigert.

Wir wollen unsere Perspektive jetzt wieder stärker auf ChatGPT eingrenzen und uns mit den Risiken beschäftigen, die hieraus entstehen können. Der Chatbot hat sich als fast allwissender Ratgeber längst bewährt, aber nicht alle Fragen sind ethisch vertretbar und dienen dem Wohl der Menschheit. Utopie und Dystopie sind oft nur eine einzige Frage voneinander entfernt: GPT-4 kann leckere Rezepte für das Abendessen vorschlagen oder tödliche Chemikalien mixen. Es kann Erste Hilfe leisten oder Menschen dazu ermutigen, sich und andere zu verletzen. Es konnte sogar Tipps geben, wie man mit einem Budget von nur einem Dollar Hunderte Menschen töten kann! Gott sei Dank verhindern mittlerweile immer wirksamere Filter bei der Ein- und Ausgabe solche unerwünschten Fragen. Diese werden sofort zurückgewiesen oder, wenn man es doch geschafft haben sollte, die Filter mit geschickten Fragen zu überlisten, werden die unerwünschten Antworten notfalls mitten im Satz abgebrochen.

Eine gute Kontrolle über die KI ergibt sich aus der Tatsache, dass diese online bereitgestellt wird, jede Frage und jede Antwort durch das öffentliche Internet läuft und also kontrolliert werden könnte. Allerdings kann man GPT und andere Modelle auch bereits lokal betreiben, ein halbwegs leistungsstarkes Notebook ist dafür ausreichend. Jenseits aller Kontrolle können lokal betriebene Systeme beliebig manipuliert und infiltriert werden, bis sie jede Art von Bullshit oder Propaganda ausspucken. Man hat an Microsoft Tay gesehen, wie schnell das gehen kann und welche Aussagen ein solchermaßen manipuliertes System plötzlich macht. *Political Correctness* kann ich einem Sprachmodell nicht nur systematisch antrainieren, man kann es auch genauso systematisch wieder entfernen. Dann mit Propaganda und Verschwörungstheorien

füttern und für jedes menschenverachtende und rassistische Statement belohnen. Auf diese Weise manipuliert, kann ich es zum Maschinengewehr für Desinformation machen und im Dauerfeuermodus auf viele Social-Media-Kanäle lenken. Natürlich haben wir all das schon von den Trollarmeen des Kremls gesehen, was aber, wenn Hunderte politische Wirrköpfe aus allen Ecken der Welt mit der hundertfachen Feuerkraft plötzlich die gesamte internationale Medienlandschaft fluten?

10.5Kann man ChatGPT auch offline nutzen?

Ja, das geht überraschend einfach, und damit entfallen viele Kontrollmöglichkeiten. Oder andersherum gesagt: Dann lässt sich die missbräuchliche Verwendung des Chatbots nicht mehr kontrollieren.

GPT4All ist ein Klon von ChatGPT und als Open-Source-Software erhältlich. Es kann schnell und einfach mit einem Windows-Installer oder auf macOS bzw. Linux lokal eingerichtet und genutzt werden. Nutzer können einen ChatGPT-ähnlichen Service im eigenen Netzwerk verwenden, ohne dass die Anfragen noch über das Internet gesendet werden müssen. Dafür muss man nicht mal in teure Hardware investieren, ein halbwegs aktuelles Notebook scheint ausreichend. Die Anforderungen von GPT4All an den Speicherplatz sind gering: Die Basisinstallation belegt rund 170 MB, die Sprachmodelle lassen sich damit auswählen und herunterladen. Diese beanspruchen dann natürlich deutlich mehr Speicherplatz, das KI-Modell **Snoozy** wird mit rund 8 GB Speicherplatz angegeben. **FreedomGPT** und **AutoGPT** sind Alternativen zu GPT4All und weitere Lösungen, um KI-Sprachmodelle lokal zu nutzen. AutoGPT beschränkt sich nicht darauf, nur als Dialogpartner zu Diensten zu sein, es kann selbstständig Aktionen ausführen und sogar ganze Projekte erstellen. AutoGPT bietet sich also an, Aufgaben mit einer lokalen KI zu automatisieren.

Die Risiken lokal betriebener Sprachmodelle liegen auf der Hand: Keine Kontrolle von Eingabe oder Ausgabe, Bugs in der Sicherung durch Filter können systematisch ausgenutzt und nicht mehr nachträglich durch Updates geschlossen werden. Mit sehr großer Wahrscheinlichkeit werden sich auch Wege finden lassen, bestehende KIs nachzutrainieren oder ganz neu anzulernen. Mit Texten, die Sie und ich nicht lesen wollen und den kenianischen Clickworkern noch heute schwer im Magen liegen. Statt diese Inhalte dann als *schlecht* und *unerwünscht* zu taggen, kann ich genau das Gegenteil tun und Diskriminierung, Verleumdung und Hassrede zum guten und erwünschten Standard erklären. Ein neuronales Netz kann Gut und Böse nicht unterscheiden und wird jede Art von

Inhalten einfach stur Heil reproduzieren.

Sie könnten einwenden, dass Diskriminierung, Verleumdung und Hassrede schon jetzt weit verbreitet sind, und hätten damit auch Recht. Wenn nun ein Heer von KIs systematisch als Soldaten für Propaganda und alle absurden Denkrichtungen eingesetzt wird, ergibt das auf Dauer eine ganz neue Qualität der Information im öffentlichen Raum und ein wahres Trommelfeuer der Desinformation.

10.6Desinformation aufdecken und bekämpfen

Im Folgenden wollen wir uns mit der Publikation *Desinformation aufdecken und bekämpfen* beschäftigen, die sich als »interdisziplinärer Ansatz gegen Desinformationskampagnen und für Meinungspluralität« versteht und in der Reihe *Schriften zum Medien- und Informationsrecht* erschienen ist. Die Herausgeber Steinebach, Bader et al. beschäftigen sich mit der Wirkung von Desinformation und vor allem mit digitaler Desinformation. Das postfaktische Zeitalter beginnt laut dieser Schrift im US-Präsidentschaftswahlkampf 2016, wo zahlreiche Falschmeldungen im Internet zirkulierten. Das Ergebnis kennen wir alle: Es wurde ein Kandidat gewählt, der immer wieder mit falschen Zahlen und unwahren Behauptungen Wahlkampf gemacht hatte. Und (Anmerkung des Verfassers) acht Jahre später könnte sich dasselbe nochmals wiederholen.

Zwar sind Reichweite und Wirkung digitaler Desinformation in Deutschland noch vergleichsweise gering, aber auch schon erschreckend hoch. Als Beleg dient ein Vorfall am Rande eines Stadtfests in Chemnitz im August 2018, wo während einer gewalttätigen Auseinandersetzung ein 35-jähriger Mann getötet und zwei Freunde des Mannes verletzt wurden. Als Täter wurden zwei Männer aus Syrien und dem Irak bezichtigt, der Getötete hätte angeblich eine deutsche Frau vor einem sexuellen Übergriff gerettet. Die wilden Gerüchte und voreiligen Interpretationen verbreiteten sich in Windeseile über soziale Netzwerke, es gab Aufrufe zu einem Trauermarsch und tatsächlich eine Demonstration von *besorgten Bürgern*, die von allen möglichen politischen Gruppierungen zur Verbreitung der eigenen Narrative genutzt wurden. Die öffentliche Wahrnehmung des Vorfalls hatte sich völlig von der Wahrheit entkoppelt. Heute könnte man weit weniger gravierende Ereignisse mit dem automatischen Trommelfeuer einer KI gezielt zu ganz anderen Wellen der Entrüstung hochkochen. Mit den leider absehbaren Folgen, die solche Vorkommnisse auf den Ausgang von Wahlen in den betroffenen Regionen oder auf Bundesebene haben

würden.

Die oben genannte Schrift (übrigens gefördert vom Bundesministerium für Bildung und Forschung) stammt zwar noch aus der Zeit vor ChatGPT (1. Auflage 2020), erkennt aber bereits die Gefahren der Social Networks für die Verbreitung von Desinformation und nennt begünstigende Faktoren wie ein hohes Tempo, Reziprozität, niedrige Kosten, Anonymität, Massenverbreitung, Passgenauigkeit und Unsichtbarkeit. Sogenannte *Social Bots* würden zur Verbreitung von Fake News, gefälschten Augenzeugenberichten und User-Kommentaren eingesetzt und somit als wichtige Instrumente in Desinformations-Kampagnen gelten. Mit den neuen Möglichkeiten generativer KI lassen sich solche Social Bots im Handumdrehen zum Taschengeldpreis erstellen und betreiben.

Der wichtigste Teil der Schrift beschäftigt sich mit der Erkennung von Desinformation auf allen Ebenen, von der Erkennung sogenannter *Malicious-Bots* (wörtlich übersetzt *bösartige Bots*) bis hin zur von Erkennung von Bildmanipulationen.

10.7 ChatGPT verschärft Cyberbedrohungen

Die Risiken von politisch motivierter Desinformation sind das eine, leider haben auch Cyberkriminelle die fortschrittliche Technologie für ihre böswilligen Zwecke entdeckt und lassen sich bei mangelnden technischen Fähigkeiten von ChatGPT auf die Sprünge helfen. Es gibt zwar bestimmte Einschränkungen, um die Nutzung des Tools für gesetzeswidrige Zwecke zu unterbinden, Cyberkriminelle finden jedoch immer Wege, solche Schranken zu überwinden. Die Nutzung von ChatGPT zu betrügerischen Zwecken umfasst das Coding von Schadsoftware sowie die Verbreitung sehr wirkungsvoller Phishing-Mails und könnte dem gefürchteten Identitätsdiebstahl ganz neue Impulse geben.

Das alles ist keine Zukunftsmusik: ChatGPT wurde bereits zur Entwicklung von Schadsoftware verwendet, hat Ransomware-Angriffe mit entsprechenden Skripten unterstützt oder hat bestehenden Malware-Code weiter optimiert. ChatGPT versteht sich gar auf polymorphe Malware mit wandelbarem Code und kann damit die Effizienz von Cyberangriffen erheblich steigern. Mit ChatGPT können sogar Laien ohne besondere technische Kenntnisse funktionierenden Malware-Code erstellen. Wir müssen uns also auf ein neues Heer von Amateur-Hackern einstellen, die die Bedrohung im Internet vervielfachen könnten. Phishing-Mails konnte man in der Vergangenheit oft an holpriger

Sprache und schlechten Übersetzungen erkennen, damit ist es nun endgültig vorbei. Cyberkriminelle und ChatGPT sind schon jetzt ziemlich beste Freunde.

Und Vorsicht: Warten Sie, bis Sie Ihr Ehepartner oder eine andere vertraute Person anruft und wegen einer unerwarteten Notlage dringend um Geld bittet – wenn die Stimme mit einer KI gefälscht wurde, haben Sie keine Chance und sind Ihr Geld schon so gut wie los! Im beruflichen Kontext müssen neue Standards für IT-Sicherheit und in der Schulung von Mitarbeitern her. Die Methoden zur Erkennung und aktiven Verteidigung gegen Cyberattacken funktionieren nicht mehr. Fehler in Grammatik und Rechtschreibung oder merkwürdige Formulierungen suchen Sie künftig vergebens in Phishing-Mails.

11Himmlische Offenbarung oder »Trojanische Bombe«?

Kann man das bisher Gesagte im Sinne eines Schlusswortes sinnvoll zusammenfassen? Welche der vielen Prognosen und Voraussagen wird denn letztlich zu unserer neuen Realität? Darauf können wir natürlich auch keine Antwort geben, aber das mit der Zusammenfassung scheint lösbar. Wir wollen hier nochmals die beiden Denkrichtungen in ihren Extrempositionen einander gegenüberstellen: Chancen und Risiken, Licht und Schatten, Utopie und Dystopie. Wir wollen aber keine weiteren Prognosen aufstellen, sondern die Entwürfe und Gedankenspiele für sich sprechen lassen.

Zu den größten Pessimisten gehört sicherlich Manfred Schneider, emeritierter Professor für deutsche Literaturwissenschaft an der Ruhr-Universität Bochum, von dem auch der Ausdruck »Trojanische Bombe« stammt. In einem Gastkommentar der Neuen Zürcher Zeitung[10] warnt er, dass generative KI »kulturelle und wissenschaftliche Verheerungen auslösen wird«. Wortmächtig und sprachgewaltig donnert er über die neuen technischen Errungenschaften hinweg, beklagt die »zermalmende Macht des KI-Konformismus« und verurteilt die Arbeitsweise der Tools mit Sätzen wie »Dem Prinzip des Neuen wird der Garaus gemacht.« Die dreiste Gleichgültigkeit, mit der die Macher der KI ignorierten, welche Rechte an Texten, Bildern und Fotos hängen, ist ebenso Thema wie die Tatsache, dass texterzeugende Maschinen die »Grundsätze der Wahrheit« nicht kennen und dass es nicht möglich ist, sie der »politischen Verantwortung zu verpflichten«. An die Stelle von Verlegern, die sich der kritischen Urteilsbildung verpflichtet fühlten, treten nun Medienmanager von Facebook, Google, Amazon und Co und »läuten dem Journalismus das Totenglöcklein«.

Ein Kulturkritiker der alten Schule, der einen strengen Blick auf die digitale Welt hat und sich darin nur schwer zurechtfindet, weil sich vieles davon nicht ins bewährte Weltbild einfügen lässt, der aber natürlich auch jedes Recht hat, negative Potenziale zu benennen und vor einem umfassenden Kompetenzverlust der Gesellschaft in drastischen Bildern zu warnen. Da braucht es viel Zuversicht und Optimismus, dem klugen und umfassend gebildeten Professor zu widersprechen oder zumindest ein paar eigene Ideen entgegenzustellen:

ChatGPT kann man auch als himmlische Offenbarung begreifen und begrüßen. Man kann es zu Recht als Motor für alle sehen, die bisher unterprivilegiert waren oder an ihrem Bildungsort überfordert: Jeder

noch so anspruchsvolle fremdsprachige Text kann in die eigene Sprache übersetzt werden, und zwar genau auf dem Niveau, das dem Leser entspricht. Und ein jugendlicher Migrant, der in der 12. Klasse des Gymnasiums die Monologe von Leicester aus Schillers Maria Stuart nicht versteht, kann sich diese von ChatGPT zuerst übersetzen und dann in der deutschen Version vereinfachen lassen. Dann hat er zwar die schillersche Poesie nicht verinnerlicht, aber zumindest so weit verstanden, dass er am Unterricht teilhaben kann. Welch eine Chance! Das wäre endlich Bildungsgerechtigkeit, lieber Herr Schneider!

Bleiben wir noch beim Thema *Lernen*, hier hat ChatGPT ein enormes Potenzial und könnte tatsächlich unsere Bildungslandschaft auf den Kopf stellen: Durch ChatGPT wird bald echtes personalisiertes Lernen möglich. Langweilige Unterrichtspausen, die für die schneller denkenden Schüler entstehen, weil die anderen den neuen Stoff nicht so rasch verstanden haben, gehören bald der Vergangenheit an. Denn der Chatbot weiß ganz schnell, was ich schon beherrsche, und kann mir an mein Niveau angepasste, interessante Lernmöglichkeiten entwerfen. Dadurch wird wertvolle Unterrichtszeit nicht mehr verschwendet, und niemand muss weiter Löcher in die Luft starren. Umgekehrt – und das ist vielleicht noch viel wichtiger – können schwächere Schüler passgenau und in exakt den Bereichen gefördert werden, wo sie Unterstützung brauchen.

Im Erwachsenenalter haben viele Menschen ein unüberwindbares Problem, das sie (gerade) niemandem anvertrauen können und wollen. Aber sie bräuchten dringend Hilfe! ChatGPT weiß vielleicht ersten Rat, ist diskret und kann ohne Scheu befragt werden. Natürlich hat die Maschine keine therapeutischen Qualitäten und ersetzt keinen Experten. Aber wenn man viele Monate Wartezeit überbrücken muss, bis man einen Therapeuten sprechen kann, mag der Bot einstweilen einen *Dialogpartner* darstellen. Das kostet sicherlich erst einmal Überwindung, aber das *Gespräch*, das sich hier entspinnt, kann helfen, die Zeit bis zum ersten Treffen mit dem realen Therapeuten durchzustehen. Und schließlich kann man ChatGPT auch als Erziehungsratgeber ohne ideologischen Rahmen heranziehen: Wie reagiert man am besten auf Trotz? Wie erziehe ich mein Kind zum Teilen? Warum schläft mein Kind nicht durch? Was tun, wenn mein Kind von Mitschülern gemobbt wird? An Antworten wird es niemals fehlen.

Man mag Manfred Schneider zustimmen, dass die Prinzipien der Wahrheit von generativen Technologien ignoriert werden. Aber im Alltag der meisten Menschen geht es um ganz banale Tätigkeiten, deren Vereinfachung ihr Leben bereichern könnte. Tabellen erstellen, Statistiken auswerten, Gliederungen anlegen, Standardmails versenden – wer hat davon schon einen Benefit? All das sind langweilige Vorgänge,

die man ohne Verlust an Kultur die Maschine erledigen lassen kann. Auf der anderen Seite gewinnt man dadurch Effizienz, und im Endeffekt wertvolle Lebenszeit, die man mit essenziellen Dingen füllen kann.

Niemand scheut sich, einen Taschenrechner zu Hilfe zu nehmen, auch um einfache mathematische Aufgaben zu lösen. Sind wir deshalb alle dümmer geworden? Nur weil wir die kleine Maschine für uns rechnen lassen? Das würde sicher niemand ernsthaft behaupten. Warum also sollten wir nicht die KI an sprachlichen Aufgaben beteiligen, sie Vorschläge für Gliederungen machen lassen? Sie nach Lösungsansätzen fragen? Nach Einleitungen für Vorträge oder Festreden?

Wir, die Autoren des Buches, gehen davon aus, dass wir auch nach monatelanger Beschäftigung mit dem Werkzeug noch lange nicht alle Potenziale entdeckt und die Möglichkeiten noch nicht annähernd ausgereizt haben. Wer ernsthaft dranbleibt, wird weitere Stärken für sich entdecken. Mein persönlicher Favorit ist die Möglichkeit zu *Fragen ohne Limit*. Wenn meine Umgebung längst die Geduld mit meiner penetranten Fragerei verloren hat, ich einen Text immer und immer wieder verändern will, eine philosophische Idee einfach nicht kapiere oder noch mehr Beispiele für ein beliebiges Problem lesen will – mein KI-Assistent antwortet immer freundlich, geduldig und vor allem klug:

```
Natürlich, hier sind einige Vorschläge …
```

Das KI-Betthupferl

Das Betthupferl ist eine der ältesten Radiosendungen des Bayerischen Rundfunks und feiert dieser Tage Geburtstag. Seit 70 Jahren gibt es fantastische Gutenachtgeschichten für Kinder, die am Ende eines langen Tages mit vielen neuen Eindrücken im Kopf noch nicht einschlafen können. Eine solche fantastische Geschichte soll auch dieses Buch beschließen, entstammt aber nicht der Feder eines realen Autors oder Journalisten. Wir haben diese Aufgabe ChatGPT (4.0) übergeben und die letzte Version ohne Eingriff oder Korrekturen in ganzer Länge übernommen. Bevor es so weit war, gab es eine ganze Reihe von Prompts, ein langes Hin und Her mit immer neuen Anforderungen und Korrekturen. Ausgangspunkt war eine simple Grundidee mit ein paar Details – hier ist das Ergebnis:

Worte der Vergangenheit

In einer nicht allzu fernen Zukunft, in der kreative Künste und wissenschaftliches Schreiben von generativen KI-Modellen übernommen wurden, verloren die Menschen allmählich ihre Fähigkeit und das Wissen, selbst komplexe Texte zu verfassen. Ein Netz aus Algorithmen erschafft, lernt und kommuniziert miteinander, indem es die Worte und Konzepte in einer endlosen Schleife zirkulieren lässt, ohne menschliches Zutun.

In dieser Welt lebt Ada, eine junge Linguistin mit einer seltenen Besessenheit für alte Sprachen und Texte, welche die einmalige Schönheit und Komplexität menschlicher Kommunikation zu schätzen weiß. Ihre Persönlichkeit ist scharf kontrastiert mit ihrer Umgebung – einer Gesellschaft, die bereitwillig das Denken an Maschinen ausgelagert hat. Ihre attraktive Erscheinung und lebhafte Ausstrahlung bleiben oft unterbewertet, da die Gesellschaft ihre intellektuellen Fähigkeiten und ihre subtile Rebellion gegen das vorherrschende System unterschätzt.

Dem gegenüber steht Dr. Elion Virtanen, ein hoch angesehener Ingenieur und Architekt des gegenwärtigen KI-Systems, der in den generativen Algorithmen die Krone menschlicher Innovation und Entwicklung sieht. Für ihn repräsentieren die Algorithmen die Verwirklichung eines Traumes, bei dem menschliches Versagen und Irrationalität durch maschinelle Präzision und Effizienz ersetzt werden.

Während die KI-Modelle weiterhin in einem zunehmend schnelleren Zyklus aus Selbstlernen und Textgenerierung agieren, bemerkt Ada eine zunehmende Degradation der Textqualität, ein langsames

Eindringen eines »geistigen Reibungsfaktors«, der die Texte allmählich zu unlogischen, bedeutungslosen Wortketten degradiert. In einer Welt, in der sich niemand mehr an die Kunst des Schreibens erinnert, beginnt Ada, heimlich ein Netzwerk Gleichgesinnter zu formen, welches sowohl die Kunst des Schreibens als auch kritische Denkfähigkeiten wiederbelebt. Sie beginnt, den mechanischen, gefühllosen Texten der KI mit poetischen, emotionalen Worten zu begegnen, die aus der menschlichen Seele entspringen. Es beginnt eine Bewegung, die die Bedeutung hinter den Worten wiederherstellt und der mechanischen Kälte des algorithmischen Outputs Wärme und Tiefe entgegensetzt.
Dr. Virtanen, der Adas Bewegung als Bedrohung für seine utopische Vision einer fehlerfreien, durch Maschinen kontrollierten Gesellschaft sieht, wird zu ihrem Antagonisten. Er glaubt, dass Adas Rebellion das gesamte System stürzen und die Gesellschaft ins Chaos stürzen könnte. Eine Jagd beginnt, bei der Ada und ihre Bewegung gegen Virtanen und sein technokratisches Regime stehen, eine Konfrontation zwischen maschinen-generierter Kälte und authentischer menschlicher Emotionalität.
Die Spannung steigt, als Ada und ihre Bewegung das KI-System infiltrieren, um eine Nachricht an die gesamte Gesellschaft zu senden – eine Botschaft, die an die verschüttete, tiefe Verbindung zur menschlichen Kreativität und Emotionalität appelliert. Während Dr. Virtanen sich bemüht, die Kontrolle wiederzugewinnen und die Rebellion zu unterdrücken, spürt er, vielleicht zum ersten Mal, eine Regung emotionaler Verwirrung, als die Worte Adas ihn erreichen und er beginnt, die erste Regung von Zweifel an seiner überzeugten Vision zu empfinden.
In der entscheidenden Konfrontation zwischen Ada und Virtanen entfaltet sich eine unerwartete Wendung, als Adas Worte eine Saite in Virtanen berühren, der plötzlich mit der Tiefe und Wärme konfrontiert wird, die die menschliche Sprache bereithalten kann. Er wird mit dem Konflikt zwischen seiner Vision einer klinisch sauberen, effizienten Welt und der unwiderstehlichen Schönheit und Komplexität der menschlichen Kreativität und Ausdruckskraft konfrontiert.
Der Roman endet mit einer unsicheren Zukunft, bei der die Tore zur menschlichen Kreativität und Expression wieder geöffnet sind, aber auch mit einer Gesellschaft, die lernen muss, sich von der Abhängigkeit von den Maschinen zu lösen und den Wert menschlichen Denkens und Fühlens wieder zu schätzen. Ada bleibt ein Symbol für die beharrliche Flamme menschlicher Kreativität und Empathie in einer kalten, mechanisierten Welt.
Der Plot lässt Raum für Reflexion über die Rolle der Technologie, menschlicher Kreativität, Emotionalität, und der beständigen Suche nach Bedeutung in einer zunehmend von Algorithmen

```
gesteuerten Welt.
```



KI-generiertes Filmplakat, erstellt mit neuroflash[11]

Quellenverzeichnis

Kapitel 5

Manuela Lenzen, Künstliche Intelligenz. Fakten, Chancen, Risiken, München 2020.

Richard David Precht, Künstliche Intelligenz und der Sinn des Lebens, München 2020.

https://www.business-humanrights.org

https://netzpolitik.org

https://www.deutschlandfunk.de/ki-kuenstliche-intelligenz-chatbot-chatgbt-100.html

Rudi Novotny/Jakob von Lindern, Alles unter Kontrolle … in: DIE ZEIT, no. 16, 13. 04. 2023, S. 31 ff.

Ann-Kathrin Nezik, Fürchten sich jetzt sogar die Erfinder?, Interview mit Yoshua Bengio in: DIE ZEIT, Nr. 16, 13. 04. 23, S. 33

Ulrich Schnabel, Unsere neue Denkaufgabe, in: DIE ZEIT, no. 21, 17. 05.23, S. 31

Harald Lesch, Marco Smolla & Hannah Bast in: Terra X Lesch und Co., #ChatGPT & Co – wie intelligent ist KI? (*https://www.youtube.com/watch?v=0fq1jlWcUv8*)

Kapitel 7

https://www.ardaudiothek.de/episode/der-ki-podcast/werden-wir-alle-arbeitslos/ard/94637888

Manuela Lenzen, Künstliche Intelligenz. Fakten, Chancen, Risiken, München 2020.

DIE ZEIT, Nr. 16, 13. 04. 2023, S. 33 (Interview mit Jens Südekum: Stiehlt uns die KI die Jobs?)

Handelsblatt Disrupt – Der Podcast über Disruption und die Zukunft der Wirtschaft (*https://www.youtube.com/watch?v=kNZQW3LA3ks*)

Kapitel 8

https://www.zeit.de/digital/2023-04/chatgpt-kuenstliche-intelligenz-forschung?utm_referrer=https%3A%2F%2Fwww.google.com%2F

Sascha Lobo in Augsburger Allgemeine Zeitung, 24.04, 2023

https://www.deutschlandfunkkultur.de/kommentar-chat-gpt-ki-100.html

Anmerkungen

1 *https://businessolution.org/de/gpt-3-statistics/*
2 Precht, Künstliche Intelligenz, S. 125
3 Precht, Künstliche Intelligenz, S. 120
4 DIE ZEIT, 13. 04. 23, S. 32
5 *https://www.bildunginbayern.de/weiterfuehrende-schule/kischool/*
6 *https://www.netzpiloten.de/4-neue-berufe-durch-kuenstliche-intelligenz/*
7 *https://www.spiegel.de/wissenschaft/mensch/maschinen-die-geklaute-buecher-lesen-christian-stoecker-kolumne-a-52041df1-6305-4969-a3b9-d781b22ed9e6*
8 *futureoflife.org/open-letter/pause-giant-ai-experiments/*
9 Jerry Kaplan: Künstliche Intelligenz (S. 143)
10 *https://www.nzz.ch/meinung/die-zermalmende-macht-des-ki-konformismus-chat-gpt-ist-eine-trojanische-bombe-ld.1730998*
11 *https://neuroflash.com/*